

GenCore version 4.5

Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: September 23, 2002, 21:12:04 , Search time 64.71 Seconds  
(without alignments)  
1208 406 Million cell updates/sec

Title: US-09-733-764-2

Perfect score: 3868

Sequence: 1 MDSIASLVICGVSLISGIV M-FAIHHYHCKSLISLSPK 704

Scoring table:

BLOSUM62

Gap: 10 0 Gap: ext 0 5

Searched: 747574 seqs, 11673796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0

Maximum DB seq length: 20000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database			
A_GenSeq_032802.*			
1:	/SIDSL/qcgdata/geneseq/geneseq-emb1/AA1980.DAT.*		
2:	/SIDSL/qcgdata/geneseq/geneseq-emb1/AA1981.DAT.*		
3:	/SIDSL/qcgdata/geneseq/geneseq-emb1/AA1982.DAT.*		
4:	/SIDSL/qcgdata/geneseq/geneseq-emb1/AA1983.DAT.*		
5:	/SIDSL/qcgdata/geneseq/geneseq-emb1/AA1984.DAT.*		
6:	/SIDSL/qcgdata/geneseq/geneseq-emb1/AA1985.DAT.*		
7:	/SIDSL/qcgdata/geneseq/geneseq-emb1/AA1986.DAT.*		
8:	/SIDSL/qcgdata/geneseq/geneseq-emb1/AA1987.DAT.*		
9:	/SIDSL/qcgdata/geneseq/geneseq-emb1/AA1988.DAT.*		
10:	/SIDSL/qcgdata/geneseq/geneseq-emb1/AA1989.DAT.*		
11:	/SIDSL/qcgdata/geneseq/geneseq-emb1/AA1990.DAT.*		
12:	/SIDSL/qcgdata/geneseq/geneseq-emb1/AA1991.DAT.*		
13:	/SIDSL/qcgdata/geneseq/geneseq-emb1/AA1992.DAT.*		
14:	/SIDSL/qcgdata/geneseq/geneseq-emb1/AA1993.DAT.*		
15:	/SIDSL/qcgdata/geneseq/geneseq-emb1/AA1994.DAT.*		
16:	/SIDSL/qcgdata/geneseq/geneseq-emb1/AA1995.DAT.*		
17:	/SIDSL/qcgdata/geneseq/geneseq-emb1/AA1996.DAT.*		
18:	/SIDSL/qcgdata/geneseq/geneseq-emb1/AA1997.DAT.*		
19:	/SIDSL/qcgdata/geneseq/geneseq-emb1/AA1998.DAT.*		
20:	/SIDSL/qcgdata/geneseq/geneseq-emb1/AA1999.DAT.*		
21:	/SIDSL/qcgdata/geneseq/geneseq-emb1/AA2000.DAT.*		
22:	/SIDSL/qcgdata/geneseq/geneseq-emb1/AA2001.DAT.*		

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	length DB	ID	Description
1	3868	100.0	704	22	AAH31165
2	3721.5	96.0	477	22	AAH31164
3	2623	67.8	1124	15	AAH45440
4	2623	67.8	1124	16	AAH73953
5	2623	67.8	1124	20	AAH30318
6	2623	67.8	1124	22	AAH65945
7	2432.5	62.9	1123	16	AAH73951
8	2432.5	62.9	1123	16	AAH67391
9	2437.5	62.8	1128	21	AAH50046
10	2427.5	62.8	1122	21	AAH59048
11	2413.5	62.4	1122	15	AAH46427

12	2412	62.4	1117	15	AAH53146	Mouse tyrosine kin
13	2388.5	61.8	1101	20	AAH84160	Rat orphan tyrosin
14	1401.5	36.2	764	21	AAH68949	Fusion protein of
15	1354.5	35.0	963	19	AAH70540	Integrin beta-1 ch
16	1304	33.7	452	22	AAH92591	Fil1 receptor fusi
17	1304	33.7	557	22	AAH97590	Fil1 receptor fusi
18	1300	33.6	462	22	AAH97592	Fil1 receptor fusi
19	1300	33.6	518	15	AAH51803	Sequence of a reco
20	1300	33.6	518	22	AAH51803	Sequence of a reco
21	1300	33.6	518	22	AAH51803	Sequence of a reco
22	1300	33.6	518	22	AAH51803	Sequence of a reco
23	1300	33.6	518	22	AAH51803	Sequence of a reco
24	1293.5	33.4	542	22	AAH97597	Fil1 receptor fusi
25	1290	33.4	497	21	AAH97172	Human FGF-R1 Extra
26	1290	33.4	525	21	AAH97171	Human FGF-R1 Extra
27	1290	33.4	622	21	AAH97170	Human FGF-R1 Extra
28	1289	33.3	474	22	AAH14051	Human FGF-R1 Extra
29	1289	33.3	528	22	AAH14051	Human FGF-R1 Extra
30	1287.5	33.3	465	22	AAH14053	Human FGF-R1 Extra
31	1285.5	33.2	631	10	AAH91009	Sequence encoded b
32	1285.5	33.2	631	21	AAH19598	Sequence encoded b
33	1285.5	33.2	631	21	AAH19598	Sequence encoded b
34	1285.5	33.2	631	21	AAH19598	Sequence encoded b
35	1284.5	33.2	497	22	AAH97597	Fil1 receptor fusi
36	1284.5	33.2	407	22	AAH86899	Human osteoprotege
37	1283	33.2	533	22	AAH13052	Human osteoprotege
38	1282.5	33.2	413	22	AAH72918	Human osteoprotege
39	1282.5	33.2	413	22	AAH80900	Human osteoprotege
40	1282	33.1	477	20	AAH90307	Human osteoprotege
41	1281	33.1	485	13	AAH24016	Human osteoprotege
42	1279.5	33.1	535	22	AAH13056	Human osteoprotege
43	1278.5	33.1	396	18	AAH18574	Human osteoprotege
44	1278.5	33.1	406	18	AAH18574	Human osteoprotege
45	1278.5	33.1	400	21	AAH15123	Human osteoprotege

ALIGNMENTS

RESULT 1

AAH31165	AAH31165 standard; protein; 704 AA.
10	AAH31165 standard; protein; 704 AA.
XX	
AC	AAH31165;
XX	
DT	02-APP-2001 (first entry)
XX	
DE	Amino acid sequence of a Tek/Fe fusion protein.
XX	
KW	Fusion protein, receptor tyrosine kinase; Tek, FC portion;
KW	immunoglobulin-like; tyrosine phosphorylation; cellular neoplasia; cancer;
KW	inflammatory disease; arthritis; rheumatism; psoriasis; eye disorder;
KW	diabetic retinopathy; retinopathy; neovascular glaucoma; retinoblastoma;
KW	retrolental fibroplasia; rubeosis; vitreous; macular degeneration;
KW	graft neovascularisation; cancer; metastatic sarcoma; carcinoma;
KW	wound granulation.
XX	
OS	Synthetic.
OS	Homo sapiens.
XX	
PH	Key
FT	Peptide
FT	location/qualifiers
FT	1..18
FT	/note "Signal peptide"
FT	19..472
FT	/note "Tek extracellular domain fragment"
FT	473..704
FT	/note "Fc portion"
XX	
FN	W0200075323-AL
XX	
EC	14 DEC 2000.
XX	
PP	97 JUN 2000, 2303W0-0215706.

XX 07-JUN-1999; 9905-0137889.  
 XX (IMMUNEX CORP.)  
 XX Carretti DP, Borges LG, Fanslow WC;  
 XX WPI: 2001-112149/12.  
 XX New Tek polypeptides antagonist having a fragment of the Tek  
 PT extracellular domain, useful for treating mammals with a disease  
 PT mediated by angiogenesis, e.g. tumors, ocular neovascularisation or  
 XX inflammatory diseases -  
 XX Claim 17; Page 37-39; 43pp; English.  
 XX The present sequence represents a fusion protein comprising the  
 CC extracellular domain of the human receptor tyrosine kinase tek and the  
 CC Fe portion of human immunoglobulin (Ig) G1. The tek fragment lacks all  
 CC or part of the region containing fibronectin type III (FNIII) motifs, and  
 CC retains the ability to bind at least one Tek ligand. The fusion  
 CC polypeptide is an angiogenesis inhibitor, and a Tek antagonist. The  
 CC polypeptide or soluble Tek multimer, antibodies or antibody fragments are  
 CC useful for treating a mammal having a disease or condition characterized by  
 CC angiogenesis, e.g. a solid tumor or disease characterized by  
 CC ocular neovascularisation. In particular, the Tek antagonists are useful  
 CC for treating or preventing inflammatory diseases (e.g. arthritis,  
 CC rheumatism or psoriasis), certain eye disorders (e.g. diabetic  
 CC retinopathy, retinopathy of prematurity, neovascular glaucoma,  
 CC retinoblastoma, retrolental fibroplasia, rubeosis, uveitis, macular  
 CC degeneration or graft neovascularisation), cancer (e.g. metastatic  
 CC sarcomas or carcinomas), or wound granulation.  
 XX Sequence 704 AA:

Query Match 100.0%; Score 3868; DB 22; Length 704;  
 Host/Local Similarity 100.0%; Pred. No. 1,40-225;  
 Matches 704; Conservative 9; Mismatches 6; Indels 0; Gaps 0;

QY 1 MDLSASLV:GWYSLISLSTVE:VAMDLLINSLPIVSDAFTSLTCTASWPPHPHETIGRD 60  
 DB 1 mdsiaslvicgvsllsytveqamdllinslplvsdaftsbtclsgswrphlepitiqrd 60

QY 61 FREALNQHJQPLEVTVTQVREWAKKVVWVREKASKIN:AYF:EGPVGEALPIPTMMRQ 120  
 DB 61 fealnmqhdpdvleqvtrwakkvwwvrekaskingayicedrvrgealrtrmmrq 120

QY 121 QASFLPATITMTVVKGDVNI:SPKKVLKREDAVIYKNGSF:HSVPKRFVHDILLEVHLPH 180  
 DB 121 qastlpatitmtvdkqdvnlstkkvlkeedavlykngslhsvprhvpdillevlhplh 180

QY 181 AQPDMGVYSAPYIGNIFTSNFTFLIVER:EA:KWIPE:NML:IA:MNN:AV:HEH:IGEC 240  
 DB 181 aqpdgagvsaryiggniftsnftflvrrecagkwpecnhictacmngvchedtgec 240

QY 241 ICPPEPMGRTEKACELITFGRTEKRGSGGCKSVYECCLPDYGCSCATGKGLQNE 300  
 DB 241 icppqlmgteekacelnlftrtkerqsggcksvyiclpdygscatgwkqlqne 300

QY 301 ACHPQYGHVCKLRCSCNNGMCHDFQNCICSPGWQCIQCHPFGTIPMTKIVDLDPHIE 360  
 DB 301 achpatyghvcklrcscnngmchdfqncicspgwqciqchpfgtprmtkivdlpddie 360

QY 361 VNSGFENPICKASGWPITNEMLTVKPDGTVLHPKDFNHTDIESVAFTIHRLLPDSG 420  
 DB 361 vnskfnpickasgwpltnemltvkdgtvvlhpkdfnhtdiesvaftihrlppdsq 420

QY 421 VVWCVNTVAGVKEPPE:SVKVLKPLNAPNVTCT:HNFAV:INISSEPYFGEPKS:DKT 480  
 DB 421 vvwcsvntvagmvepkpnlsvkvlkplnapnvtctahntavinnissepyfgepkscdk 480

QY 481 RVCPPDPAPELLGDSVFLFPHKPEKNTLMSKTPPEV:CVVVVSHVEDPKVKNYVDCVE 540

DB 481 hccppcpapellgdsvflfppkpkdtilmsrtpevtcvvshvedpkvknwydcve 540  
 QY 541 VINAKTKPREQYNSTYRVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTIISKAKGQP 600  
 DB 541 vlnaktkpreeqynstyrsvsvltvlhqdwlngkyckvsnkalpapiektiskakgqp 600  
 QY 601 REPQVYTLPPSRSEMIKNQVSLTCLVKGKVFYSQIAVWESNCPENNYKTFPPVLDSSGS 660  
 DB 601 repqvytlppsrseemiknqvsltclvkgkvyfsqiavwesncpenynyktfppvldsdgs 660  
 QY 661 FFLYSKLTVDKSRWQGVNFCSSVMHEALHNHYTOKSLISLSPGX 704  
 DB 661 fflyskltvdkswqgvnfcscvmhealhnytkslislspgk 704

RESULT 2  
 AAB31164  
 ID AAB31164 standard; protein; 977 AA.  
 XX AAB31164:  
 XX 02-APR-2001 (first entry)  
 XX Amino acid sequence of a Tek/Fe fusion protein.  
 XX Fusion protein: receptor tyrosine kinase; Tek; Fe portion;  
 KW immunoglobulin G1; IgG1; angiogenesis; tumor; ocular neovascularisation;  
 KW inflammatory disease; arthritis; rheumatism; psoriasis; eye disorder;  
 KW diabetic retinopathy; retinopathy; neovascular glaucoma; retinoblastoma;  
 KW retrolental fibroplasia; rubeosis; uveitis; macular degeneration;  
 KW graft neovascularisation; cancer; metastatic sarcoma; carcinoma;  
 KW wound granulation.  
 XX Synthetic.  
 OS Homo sapiens.  
 XX Key Location/Qualifiers  
 FT Peptide 1..18  
 FT Protein /note= "signal peptide"  
 FT Protein 19..745  
 FT Protein /note= "tek extracellular domain"  
 FT Protein 746..977  
 FT Protein /note= "Fe portion"  
 XX W0200075323-A1.  
 XX 14-DEC-2000.  
 XX 07-JUN-2000: 2000WO-0515705.  
 XX 07-JUN-1999: 9905-0137889.  
 XX (IMMUNEX CORP.)  
 XX Carretti DP, Borges LG, Fanslow WC;  
 XX WPI: 2001-112149/12.  
 XX New Tek polypeptides antagonist having a fragment of the Tek  
 PT extracellular domain, useful for treating mammals with a disease  
 PT mediated by angiogenesis, e.g. tumors, ocular neovascularisation or  
 XX inflammatory diseases -  
 XX Claim 41; Page 34-37; 43pp; English.  
 XX The present sequence represents a fusion protein comprising the  
 CC extracellular domain of the human receptor tyrosine kinase tek and the  
 CC Fe portion of human immunoglobulin (Ig) G1. The tek fragment lacks all  
 CC or part of the region containing fibronectin type III (FNIII) motifs, and  
 CC retains the ability to bind at least one tek ligand. The fusion  
 CC polypeptide is an angiogenesis inhibitor, and a tek antagonist. The  
 CC polypeptide or soluble tek multimer, antibodies or antibody fragments are

CC useful for treating a mammal having a disease or condition mediated by  
 CC angiogenesis, e.g. a solid tumour a condition or disease characterized by  
 CC ocular neovascularization. In particular, the Tek antagonists are useful  
 CC for treating or preventing inflammatory diseases (e.g. arthritis,  
 CC rheumatism or psoriasis), certain eye disorders (e.g. diabetes,  
 CC retinopathy, retinopathy of prematurity, neovascular glaucoma,  
 CC retinoblastoma, retrolental fibroplasia), rubeosis, uveitis, macular  
 CC degeneration or graft neovascularization), cancer (e.g. metastatic  
 CC sarcomas or carcinomas), or wound granulation  
 XX  
 S0 Sequence 977 AA:

Query Match 26.2%, Score 3721.5, DB 22; Length 977;  
 Best Local Similarity 72.1%; Pred No. 3.4e 216;  
 Matches 704; Conservative 0; Mismatches 0; Indels 273; Gaps 1;

QY 1 MDSIASLVIGVSLISQIVFEGAMDIUNSLPLVSIAPISUICIASVPPHPPHPTIGPD 60  
 Db 1 mdsiaslvicqslisgrvvgamdilinslplsdaetsltetiasgrphcptitgrd 60

QY 61 FEALMNHQJQLEVTQDVTEWAKKVVWKKKASKINGAYCEGRVRGREATIPTMKMRQ 120  
 Db 61 fealmnhqjqlvtevtqdvtewakkvwwkkaskkingayfecgrvrgcaitirtemkmrq 120

QY 121 QASFLPALHFWIVQKGLVNNISKKVLKKEEAVIYKNSFTHSVPRKRVVPLLEVHLPH 180  
 Db 121 qasflpalhtwivqkglvnniskkvllkeedavlyknsfthsvprhevpllevehlph 180

QY 181 AOPDQAGVYSARYIGGNLPTISATRIIVRRCBAOKGPCNHLCTACMNGVCHEDYGCRC 240  
 Db 181 aopdaqvysaryiggnlptisatrlivrrcbaokgpcnchlctacmngvchedygcrc 240

QY 241 ICPGFMGRKTEKATLHTRGRPTKERASQBERKSYVPLECPYASVATSWKGLQVNE 300  
 Db 241 icpgfmgrtceakclhtfgrtcekeresygqeksyvflcpdygscatgwkylqne 300

QY 301 ACHPCFVGPCKTICSNNGFMCTHRCQGLSLGWGLQCEKELIKMTIKIVLPDHLK 360  
 Db 301 achpcfvgpckircsnngfmcthrqglslgwglqcekerlikmtikivlpdhlk 360

QY 361 VNSCKFNPICKASQWLPITNFMETLVKPDCTVLRPKDFNHTDHFVAIPTHRIEPPDSG 420  
 Db 361 vnsckfnpickasqwpitnfmeltvkdptvlrpkdfnhtdhfsvaiftlhrileppdsq 420

QY 421 VWGCSVNTVAQWVEFFENISVKVLKPELAPNVIETSHNFVINSSEHYEG 472  
 Db 421 vwgcsvntvawvveffenisvkvlpelapnvnietshnfvinsseshyeg 472

QY 473 473 472

QY 481 LLYKPVHYQWQHIGVQVNEIVILNyleprteylevqlvrrqeqeqhqpvrifltas 540  
 QY 473 473 472

QY 541 IGLPPPGILILPKSQTILTLWQPIFPSSDIFYVWVTSVQSKDPPIKVPNGTFSV 600  
 QY 473 473 472

QY 601 LHLHPGQVYVNAVNTKAGQWSEDLTATLTSLIPPAENIKISNTHSSAVISWT 660  
 QY 473 473 472

QY 661 ILDQYSISSILIRYKVGQKNEQHDVVKLNATLIQYQIKGIEPETAYQVDIRAENNAQS 720  
 QY 473 473 472

QY 721 SNAFISHELVTLPSCQAPADYGGKPKSKDKTCTEPPGAPCLIGGTSVFIIPKPKKDT 780  
 QY 508 LMSRTPPEVTVVVDVSHEDPEYKWNVYDGVGVHNAKTKPREQYNSTYRVVSVLTVLH 567  
 Db 781 lmsrtpevtcvvvdvshedpevkwnvydgvgvhnaaktkpreeqynstyrvvsvltvlh 840

QY 568 QLVNCKRKYCKKYSNKALPAPTEKIIISKAKQUPREPVVILPPSKFEMIKNOVSLICLVK 627  
 Db 841 qdwlngkeykckysnkalpapektliskakqqrpeqvylppskfemiknoqslilelvk 900

QY 628 GFYPSDTAVFWSNMGZFPENNYKTTTPPVILSDSFFFLYSKLTIVDKSPWJQJNVFSSVMIIE 587  
 Db 901 gfypsdtavfwsnmgzfpennyyktppvilsdsffflyskltivdkspwjqjnvfssvmiie 960

QY 588 ALHNHYTKQSLSLSPGK 704  
 Db 961 alhnhytkqslslspgk 977

RESULT 3  
 AAR45440  
 ID AAR45440 standard; Protein, 1124 AA.  
 XX  
 AC AAP45440;  
 XX 25 JUN-1994 (first entry)  
 XX Human orphan receptor kinase.  
 DE  
 XX  
 KW Lrk; ligands, antibodies, PCR, amplification.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT Peptide 1..18  
 FT Region 746..772 /note= "signal peptide"  
 FT Region 211..340 /note= "transmembrane region"  
 FT Region 211..340 /note= "EGF-like repeat"  
 FT Misc-difference 44 /note= "Cys of immunoglobulin domain"  
 FT Misc-difference 102 /note= "Cys of immunoglobulin domain"  
 FT Misc-difference 19..1124 /note= "claimed fragment"  
 FT Misc-difference 19..745 /note= "claimed fragment"  
 XX  
 UN W09400469-A.  
 XX  
 ID 06-JAN-1994.  
 XX  
 XX 25-JUN-1994; 93WO-US06093.  
 XX  
 XX 26-JUN-1994; 94US-0905600.  
 XX  
 PA (IMMV ) IMMUNEX CORP.  
 XX  
 PI Ziegler SF;  
 XX  
 XX WPID: 1994-026142/03.  
 DR N-PSDB; AAQ55179.  
 XX  
 FT DNA and protein sequences for orphan receptor tyrosine kinase -  
 FT and expression vectors for produ. of recombinant protein and  
 FT antibodies specific for the protein, useful in research  
 XX  
 PS Claim 13; Fig 1; 57pp; English.  
 XX  
 CC Degenerate oligonucleotide primers based on the sequence conserved  
 CC in the kinase domain of all receptor tyrosine kinases was used for  
 CC PCR of single stranded cDNA from human placental polyA mRNA. PCR  
 CC prod. HPK-6 contained a novel sequence which was used as a probe to  
 CC isolate longer fragments from a human placental cDNA library. One  
 CC clone contained the entire coding region and was called the ork gene.  
 CC The gene prod. shown can be used as a research tool in in vitro assays  
 CC for detection of ork, its ligands or their interactions.  
 CC See also AAR45441.

```

XX
SQ Sequence 1124 AA:

Query Match 67.8%; Score 2623; DB 15; Length 1124;
Best Local Similarity 77.1%; Pred. No. 6,9e-150;
Matches 501; Conservative 23; Mismatches 44; Indels 82; Gaps 7;

QY 1 MDLSASLVLCGVSLISGIVFCAMDIILINSLPLVSDAETSLTCTASQWPPHIEPITGPD 60
Db 1 mdslaslvlcgvslisqslsqvqdamdlilinsplvsdaetsltctasqwrphiepitlgd 60
QY 61 FEAIMNHOCHPLIEVQDVIIRWAKKVVWKRKASKINGAYFCGKVRGEAIRITMKMRQ 120
Db 61 feaimnhochplievqdvirwakkvwwkrkaskingayfcgkvrgeairitmkmrq 120
QY 121 QASFLPATLIMIVDQGVNVSFKKVLIFKFFAVIYKKW:SFTHSVPRHIEVPDILEVILPH 180
Db 121 qasflpatlimgvqgvnvsfkkvllfkffavikykw:stfhsvprhievdpdilevilph 180
QY 121 qasflpatlimgvqgvnvsfkkvllfkffavikykw:stfhsvprhievdpdilevilph 180
Db 121 qasflpatlimgvqgvnvsfkkvllfkffavikykw:stfhsvprhievdpdilevilph 180
QY 181 AQPQAGVYSARYIGNLTSAFTBLIVRCEAOKWGPCNHLCTACMNGVCHEDTGE 240
Db 181 aqpqagvysaryignltsaftblivrceakwgpccnhlctacmngvchetedge 240
QY 241 ICHPCHMGPTCKACELHIFGTCRKESGGQKGSVYVCLPDPYGCSCATGWKGLQCNE 300
Db 241 icpchmgptckacelhifgtrkeresggqgksyvcldpdygcscaatgwkglcne 300
QY 301 ACHPFIYAPDKLRIS'NNGEM'GFFQK'LS'PQWGLQ'EREGIPMTPKIVDLPDIE 360
Db 301 achpfyapdklriss'nngem'gffqk'ls'pqwglq'eregipmtpkivdldpie 360
QY 361 VNSGRFNPICKASQWPLTNEPMTLVKPDGTVLHPKDFNHTDHSVAIFTHRIIPDSG 420
Db 361 vnsgrfnpickasqwpplnepmtlvkpdgtvlhpkdfnhtdhsvalfthriipdsq 420
QY 421 VVCSVNTVAGVWRKPPNLSYKVLFPKLNAPNVIDTGHNFAVINISSEPYGFHFKSCDKT 480
Db 421 vvcsvntvagvwrkppnlsvkvlfpklnapnvidtghnfnavinissepygfghfkscdkt 480
QY 481 HTCPGPAPELLGSPVFLFPFKPDTLMISKRIPEVICVVDVDSHEDPEKFNMYV 536
Db 481 htcpgpapellgspvflfppfkpdtlmiskripevicvvdvdsshedpekfennvy 536
QY 474 ---gpiakskkll-----ykpnyhycawqlqvtnclvtnlylqprteyeyelqvtr 521
Db 474 ---gpiakskkll-----ykpnyhycawqlqvtnclvtnlylqprteyeyelqvtr 521
QY 537 --DGVEVHNAKTKPREQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTI 594
Db 537 --dgvevhnaktkpreqynstyrvvsvlvtvlhqdwlngkeykckvsnkalpapiektis 594
QY 522 rgeqeqh -----ppprvrrfrrt 538
Db 522 rgeqeqh -----ppprvrrfrrt 538
QY 595 KAKGGRKRPQVYITLPPSR--EEMTKNQVSLTCLVKGFYPS---DIAVEWE 619
Db 595 kakggrkrpqvyltlppsr--eemtknqvsltclvkgyfyps---diavewe 619
QY 539 asig-----lppqqlnlpksgtlnltwqplfpsseddfyveve 579
Db 539 asig-----lppqqlnlpksgtlnltwqplfpsseddfyveve 579

RESULT 5
AAY30318
ID AAY30318 standard; protein: 1124 AA.

```



XX AAY30318;  
 XX 15-NOV-1999 (first entry)  
 XX Amino acid sequence of TEK (also known as TIE2) protein.  
 XX TEK protein; TIE2 protein; receptor tyrosine kinase; T cell response;  
 XX immune response; endothelial cell; tumor associated vasculature;  
 XX coagulation; thrombosis; cancer; anticancer vaccine.  
 XX Homo sapiens.  
 XX WO9943801-A1.  
 XX 02-SEP-1999.  
 XX 26-FEB-1999; 99WO-GR00583.  
 XX 26-FEB-1998; 98GB-0004121.  
 XX (CANC-) CANCER RES CAMPAIGN TECHNOLOGY.  
 XX Durrant LG, Hewett PW, Ramage JM, Spendlove I;  
 XX WPI; 1999-540586/45.  
 XX New peptides containing at least one epitope from Tek receptor  
 XX tyrosine kinase, used in vaccines against cancer  
 XX disclosure; Fig 1; 56pp; English.  
 XX The present sequence represents the TEK protein, also known as TIE2  
 XX protein. TEK is a receptor tyrosine kinase. TEK contains epitopes  
 XX which bind to MHC. The presentation of TEK epitopes can also stimulate  
 XX helper cell and/or cytotoxic T cell responses. The immune response is  
 XX directed against endothelial cells in the tumor-associated  
 XX vasculature and includes production of antibodies that bind to the  
 XX cells, causing coagulation and thrombosis. The immune response is  
 XX targeted to endothelial cells lining blood vessels of the tumor (these  
 XX cells overexpress Tek), so damage to even a few cells will kill many  
 XX tumor cells. These target cells are accessible to the immune response  
 XX and problems of antigenic heterogeneity, MHC loss and resistance to  
 XX apoptosis (associated with epithelial cells) are unlikely to occur in  
 XX normal endothelial cells. TEK epitopes (see AAY30320-24) are used to  
 XX generate antibodies, and for prevention and treatment of cancer.  
 XX The peptides, and recombinant DNA constructs or viral vectors that  
 XX express them, are useful as anticancer vaccines to target endothelial  
 XX cells that line blood vessels of the tumor.  
 XX Sequence 1124 AA;  
 XX  
 XX Query Match 67 89, Score 2623, DB 20; Length 1124;  
 XX Best local Similarity 77 14; Pos. No. 6, 9e-150;  
 XX Matches 501; Conservative 23; Mismatches 44; Indels 82; Caps 7;  
 XX  
 XX 1 MDSLASLVGVSLISLSTVFGAMDILINSPLVSDAETSLTICASWPPHPIITIGRD 60  
 XX 1 mdsiaslvicvslislsatvgamdilinsplvsdaetslticagwpphpiitigrd 60  
 XX  
 XX 61 FEALMNHQHPLEVTIGQVTRFAWAKVWVKREKASKINCAVFCGKVRGLAIRIKMKMQ 120  
 XX 61 fealmnhqhpelvtigqvtrfawakvkwkrekasklscgylceqvrqgairlrmkmrq 120  
 XX  
 XX 121 QASFLPATLMTVDKGNVNIISPKKVLKPNDAVYKNCSEIHVSVPHPVPHLEVLPH 180  
 XX 121 qasflpatlmtvdkgdnvnisfkvlkcpndavikngsfihsvphphvphlevlph 180  
 XX  
 XX 181 AQPDAGVYSARYLQSLFTSAFTPLIVRCFAGKWKSPENHLGTAQMNWVWTEEDIEP 240  
 XX 181 aqpdagvysarylqslftsaftplivrcfagkwwkspenhlgtamnwvwtteediep 240  
 XX  
 XX 181 aqpdagvysarylqslftsaftplivrcfagkwwkspenhlgtamnwvwtteediey 240

QY 241 ICPPGCMGRTOCKACRLHTFGRTCKRCQSGQKSVVFTFPPVQCSGATGKGIQCNPF 400  
 DB 241 icppgcmgrtoekacelhtfgrtckercsgqeqcsyviclppdygcsaatgkqlaqce 300  
 QY 301 ACHPAPFYGPFA\*KLKNS\*NNKSEM\*DKKQKGLSPGQGGIQCRRGGTPRMTPKIVDLPDHT 360  
 DB 301 achpafyfpdcklrscengemcdrfggclsggwdglqceetgipmtpkivdipdht 360  
 QY 361 VASGKFNPIIKASGWPLPTNEEMTLVKPQGVILHFKDFNHTGHFSVAIEIIRKILLFPGSS 420  
 DB 361 vnsqkfnpicasqwpplneemtlvkpqtvlhpkdfnhtghisvalitirillpdsq 420  
 QY 421 VWCVSNTVAGWVKDPNLSVKVLKPLNAPNVITGHNPAVINISSHPYFGHPKSCDKT 480  
 DB 421 vwcvsntvagwvdkpnlsvkvlkplnapnvitghnplavinissshpyfgkpscdkt 480  
 QY 481 HTCPGCPAPELGGPSVFI\*PPKPKDILMLSKTPEVTCVVVDVSHEDPEVKENWVY 536  
 DB 474 ---gfikskvll-----yknhycawghlqfnctvctgylepteyclvqlyr 521  
 QY 537 --DGVEVHNAKTKPREEQYNSTYFVVSVLTLHQLWLN:KEYKCKVSNKALPAPIKTIS 594  
 DB 522 rdeqeqeqh-----lpprqnlipksqtlititwqilpsseddiyyeve 579  
 QY 595 KAKGQRRPQGVYTLDPGR--EEMIKNQVSLTCLVKGFYPS---DIAVEWE 639  
 DB 539 aslq-----lpprqnlipksqtlititwqilpsseddiyyeve 579  
 RESULT 6  
 AAC65945  
 ID AAG65945 standard; protein; 1124 AA.  
 XX  
 XX AAG65945;  
 XX 11-FEB-2002 (first entry)  
 XX Amino acid sequence of human tie-2.  
 XX Tie-2; catalytic domain; receptor tyrosine kinase, crystalline; human;  
 XX cytosolic; vasotropic; antianemic; antiarteriosclerotic; nephrotropic;  
 XX ophthalmologic; hepatotropic; antithyroid; antiinflammatory; antiulcer;  
 XX gastrointestinal; antirheumatic; osteopathic; antiarthritic; humoral;  
 XX antipsoriatic; dermatological; immunosuppressive; antibacterial.  
 XX Homo sapiens.  
 XX Key Location/Qualifiers  
 XX Domain 802..1124  
 XX /note= "catalytic domain; specifically claimed fragment"  
 XX  
 XX WO200172778-A2.  
 XX 01 OCT 2001.  
 XX 20 MAR 2001; 2001WO-US08853.  
 XX 29-MAR-2000; 2000RS 192920P.  
 XX (BADI ) BASF AG.  
 XX Rump NJ, Arnold LD, Dixon RW, Hoeffer HW, Allen K, Hellmanna G;  
 XX WPI; 2001-648437/74.  
 XX Crystalline polypeptide useful for identifying inhibitors of a tie 2  
 XX protein as well as determining the three dimensional structure of a  
 XX polypeptide comprising the catalytic domain of a tie-2 polypeptide -  
 XX claim b; Fig 1; 24pp; English.  
 XX The invention relates to a crystalline polypeptide, comprising the  
 XX catalytic domain of a receptor tyrosine kinase Tie 2 protein. The

crystalline forms are useful for identifying inhibitors of a Tie-2 protein as well as determining the three dimensional structure of a the catalytic domain of a Tie-2 polypeptide. A Tie-2 inhibitor may be used to treat a Tie-2 dependent condition in a patient (especially a human), where the condition is characterized by excessive vascular proliferation e.g. a hyperproliferative disorder, cancer (e.g. sarcoma, osteoma, melanoma, lymphoma, and leukemia), a cardiovascular condition (e.g. atherosclerosis, ischemia, anemia, and vascular leakage disorders), an ocular condition (myopia, chronic retinal detachment, conjunctivitis, retinopathy, and macular degeneration), von Hippel Lindau disease, hemangioid, psoriasis, Paget's disease, polycystic kidney disease, fibrosis, sarcoidosis, cirrhosis, thyroiditis, Oslar-Weber-Rendu disease, chronic inflammation, synovitis, inflammatory bowel disease, Crohn's disease, rheumatoid arthritis, osteoarthritis, psoriatic arthritis, and ulcer or sepsis, especially where the disorder involves aberrant endothelial periorbital interactions. The Tie-2 inhibitor may be used to decrease fertility, and promote angiogenesis or vasculogenesis (in combination with a pro-angiogenic growth factor). The present sequence represents the human Tie-2 protein sequence.

XX Sequence 1124 AA;

Query Match 67.8%; Score 3623; DB 22; Length 1124;  
 Best Local Similarity 77.1%; Pred. No. 6.9e 150;  
 Matches 501; Conservative 23; Mismatches 44; Indels 82; Gaps 7;

QY 1 MDLSASLVGVSLTSCIVGAMDI I I I N S P I V S N A P I S I G C I A S G W R P H P I T I G R D 60  
 Db 1 mdlsaslvicvslstveqamdlilinslplvsdaelsltciasqwrphcpitlqrd 60

QY 61 FEALMHQHDPLEVTQDVTREWAKVVKWPKASKINAYGFGVRVGEATIPITMKMPQ 120  
 Db 61 fealmhqdplvtqdvtrwakkvwwkrcaskindayfcedgrvrycalirrtkmrj 120

QY 121 QASELPATLTMTVDKGDNNVLSFKKVLKEEDAVLYKNSFIHSVPHEVPDILLEVHLPH 180  
 Db 121 qaselpatltmtvdkgdnnvlsfkvllkeedavlyknsfihsvprhevpdillevlhpl 180

QY 181 AQDQAGVVSARYIGENLFI S A P I R I L V R G C A Q K W G P P C N H C I A C M N N C V C H E D I G E C 240  
 Db 181 aqdqagvvsaryiqnllsattllvrceaqkwpccnhclctacmnnqvcchediqec 240

QY 241 ICPPGPMGHTTEKALELITFGTTEKESQLEAKSYVELEPPYGSNCAIGWKGLOCNE 300  
 Db 241 icppgpmgttekalellitfgttekesqleaksyvelppdygscatgwkglocne 300

QY 301 ACHPQFYGPGRKLRPSNNNGEMTDFWGLTSPGWGLQCEPEKIPRMTPKIVLPDHPHF 360  
 Db 301 achpqfygpgrklrpsnnngemtdfwgltpgwglqcepekiprmtpkivlpdhphe 360

QY 361 VNSGKNPDKCKASWDIPIHFMPTVRYDQCVIHWKDNHEDHRSVAFTIHPILIPDSC 420  
 Db 361 vnsqknpdkckaswdipi hfmptvrydqcvihwkdnhedhsvaiftihp ilipdsq 420

QY 421 VVWGSVNTVAGVVERPENTSVKVLPEKINAPNVIITGHNFAVINLSPEYGEPEKSCDKT 480  
 Db 421 vvwsvntvagvverpentsvkvlpekinnapnviitgtnfavinlspeygepekscdkt 480

QY 481 HTCPGPCAPPELLGSPVLEFPFKPKDITLISRTPEVTGVVDVSHEDPEVKFMNVY 536  
 Db 474 ---qpksskll-----ykpshyeaqhlyqneilvlyleprleyelcvqlvr 521

QY 537 --DGVVHNNAKIKPREHNSYTRVVSVLVILHQLWLNKRYCKKVSNNKALPAPLEKTS 594  
 Db 522 rdeqqeqh-----ppqvrftt 538

QY 595 KAKGAPPEPVVTLDPSP FEMTPTNVSLTGLVKEFYES---DIAVEWE 639  
 Db 539 asig-----lpprgnlilpksqtlntwqpfpseddfyove 579

RESULT 7

AAR73951  
 ID AAR73951 standard; Protein; 1123 AA.  
 XX AAR73951;  
 AC AAR73951;  
 DT 21-JAN-1996 (first entry)  
 XX Mouse Tie-2 receptor tyrosine kinase protein.  
 DE Mouse Tie-2 receptor tyrosine kinase, DNA primer; cancer; angiogenesis;  
 XX vasculogenesis.  
 KW tie-2; receptor tyrosine kinase, DNA primer; cancer; angiogenesis;  
 NW vasculogenesis.  
 OS Mus musculus.  
 FH Key Location/Qualifiers  
 FT Cleavage site 22..23 /note= "signal peptide cleavage site"  
 FT Misc-difference 44 /note= "involved in sulphydryl bonding of Iq domain"  
 FT Misc-difference 102 /note= "involved in sulphydryl bonding of Iq domain"  
 FT Region 211..251 /note= "epidermal growth factor-like repeat"  
 FT Region 255..298 /note= "epidermal growth factor like repeat"  
 FT Region 302..340 /note= "epidermal growth factor-like repeat"  
 FT Domain 437..720 /note= "includes 3 fibronectin type-III domains"  
 FT Domain 829..921 /note= "tyrosine-kinase domain"  
 FT Domain 936..1081 /note= "tyrosine kinase domain"  
 FT Region 135..137 /note= "RGD triplet"  
 XX W09513387-AL.  
 PD 18-MAY-1995.  
 XX 12-NOV-1994; 94W0-EF03767.  
 XX 12-NOV-1993; 93US-0152552.  
 XX (PLAC ) MAX PLANCK GES FORFORDERUNG WISSENSCHAFTEN.  
 PA Risau W.  
 XX WPI: 1995-194105/25.  
 DR N-PSDB; AAQ91999.  
 XX New Tie-2 receptor tyrosine kinase and related nucleic acid and methods for detecting Tie-2 modulators for treating eg cancer, associated with angiogenesis and vasculogenesis  
 PS Claim 3; Page 45; 81pp; English.  
 XX This protein may be expressed recombinantly in a host cell. The expressed protein may be used to treat diseases, or processes, associated with angiogenesis and vasculogenesis, or cancer. Cells that express the protein are used in screening procedures and recombinant proteins can be used for affinity purification of Tie-2 ligand.  
 SO Sequence 1123 AA;

Query Match 67.8%; Score 3623; DB 22; Length 1124;  
 Best Local Similarity 77.1%; Pred. No. 6.9e 150;  
 Matches 501; Conservative 23; Mismatches 44; Indels 82; Gaps 7;  
 QY 1 MDLSASLVGVSLTSCIVGAMDI I I I N S P I V S N A P I S I G C I A S G W R P H P I T I G R D 60  
 Db 1 mdlsaslvicvslstveqamdlilinslplvsdaelsltciasqwrphcpitlqrd 60

```

Db      i mdslaqlvicaqvsiilgyvveqamdliilinslipvsdaetslctasqwnhpeliligrd 60
QY      61 FLA:MNQHQPDLVTC:VTFEFAKAKVWVKRASKINCAVFCGKVPFGAIPRIKMKMQ 120
Db      61 foalnmqhdplovtqvtrewakkkwvkraskingayfegrvrgdairirtkmqr 120
QY      121 QASFLPATLTMTVDKINNNVNSFKKVLKEEDAVIYFNISFHSVPHEVPDILEVHLPH 180
Db      121 qasflpatlntvdrgdavn:isfkv:ikeedavlykgsfihsvprhcrpdlilcvhph 180
QY      181 AQPDQACVYSARVIGGNLTPTASFTPLIVPPCPAOKWGPCNHLCTACMNGVCHRYTPC 240
Db      181 aqpdgaqvysaryigqnltsatrlvirvrcqaqkwgpcsrctctckangvchedtgc 240
QY      241 ICPGFMGRTCEKACFTIHUPCEICKRCSQCECKSVYFCLIDPYGCSGATQWKLQONE 300
Db      241 icpgfmgrtcekaeceptftrtckercsqpegcksvyfcldpygcsatgwlgqnc 300
QY      301 ACHPGFYGFCCKLRCS:NNHEM:LVKPGVGLSPWGLSCLQCEGIPRMTPKIVDLPDHE 360
Db      301 achpgfygfccklrct:nnhe:lvkpgvglspwglscldqegiprmtpkivdldphie 360
QY      361 VNSGKFNPICKASGWPPTNEEMTLVKPDTVLHPKDFNHTDHSVAFTTHRIILPDSG 420
Db      361 vnsghfnpicakagwpptneemtivkpdgtvlpqndfnytdrfsvaitvnrviipdsq 420
QY      421 VWCSVNTVAGMVKFPNISVKV:PKPLNAPNVIDTGHNFPAVINISSEPFYC 472
Db      421 vwcsvntvagmvekpfnisvkv:peplhapnvldtghnfalinssepfyfgdpikskk 480
QY      473 -----EPKS-----CDKTHCTPCPAPPELLGSP-----S 496
Db      481 lfykpvndqakylvctneiftlnyletrdyele---vglarpgqegglpgvrrfttas 538
QY      497 VFLPFPKPKDITLMISRTIP-EVTCVVVDVSHEDPEVKENWYVDGVVHNAKTKPREQYNS 555
Db      539 iglfpfpkglslpksqalntlwpilfused-----ely---veverisqlttdqgn- 589
QY      556 TYKVVSVLT-VLHQDWLNKGYCKKVSNNKALPAPIEKTISKAKGQ-PRHQVYVY----LP 609
Db      590 -ikvpqnltsvllsnlvpreqyvra-----rvntkagqewseclrawtldslilp 638
QY      610 PSKEEM-----TKNVSLTCLVKGFYPSDIAVEWESNQPENNY 648
Db      639 ppenikisnitdstamvswt-ivdgysissliirykvqgknedqh 683

RESULT      8
AA067391
ID      AAR67391 standard; Protein; 1123 AA.
XX
AC      AAR67391;
XX
XX      24-AUG-1995 (first entry)
XX      Murine tyrosine kinase receptor tie-2.
XX      Murine tyrosine kinase receptor; tie-2; hematosis diagnosis.
XX      Mus musculus.
OS
XX      Key
XX      Peptide
XX      /label= sig_peptide
XX      Protein
XX      /label= mat_protein
XX      JP06315382-A.
XX      15-NOV-1994.
XX      06-MAY-1993; 937P-0129912.
XX

```

```

PP      06-MAY-1993; 937P-0129912.
XX      (SUDA/V) SUDA T.
XX      (YAMA ) YAMANOUCHI PHARM CO LTD.
XX      WPT: 1995-032331/05.
XX      N-PSDB: AA075334.
XX      DNA coding a tie-2 receptor and a tie 2 receptor - used in the
XX      diagnosing hematosis
XX      Claim 2; Pages 13-17; 17pp; Japanese.
XX      AA075334 encodes AAR67391 the murine tyrosine kinase receptor tie-2.
XX      The cDNA is thought to participate in hematosis, and can therefore
XX      be used in the diagnosis of hematosis.
XX      Sequence 1123 AA;

Query Match      62.9%; Score 2432.5; DB 16; Length 1123.
Best Local Similarity 68.1%; Pred. No. 2.1e-138;
Matches 481; Conservative 50; Mismatches 94; Indels 81; Gaps 14;

QY      1 MDSLASLVLCGVSLLSSTVREGAMDLILINSPLVSDAETSLCTASQWHPHEPILIGRD 60
Db      1 mdsllaglvlgvslilgyvveqamdliilinslipvsdaetslctasqwnhpeligrd 60
QY      61 FLA:MNQHQPDLVTC:VTFEFAKAKVWVKRASKINCAVFCGKVPFGAIPRIKMKMQ 120
Db      61 foalnmqhdplovtqvtrewakkkwvkraskingayfegrvrgdairirtkmqr 120
QY      121 QASFLPATLTMTVDKINNNVNSFKKVLKEEDAVIYFNISFHSVPHEVPDILEVHLPH 180
Db      121 qasflpatlntvdrgdavn:isfkv:ikeedavlykgsfihsvprhcrpdlilcvhph 180
QY      181 AQPDQACVYSARVIGGNLTPTASFTPLIVPPCPAOKWGPCNHLCTACMNGVCHRYTPC 240
Db      181 aqpdgaqvysaryigqnltsatrlvirvrcqaqkwgpcsrctctckangvchedtgc 240
QY      241 ICPGFMGRTCEKACFTIHUPCEICKRCSQCECKSVYFCLIDPYGCSGATQWKLQONE 300
Db      241 icpgfmgrtcekaeceptftrtckercsqpegcksvyfcldpygcsatgwlgqnc 300
QY      301 ACHPGFYGFCCKLRCS:NNHEM:LVKPGVGLSPWGLSCLQCEGIPRMTPKIVDLPDHE 360
Db      301 achpgfygfccklrct:nnhe:lvkpgvglspwglscldqegiprmtpkivdldphie 360
QY      361 VNSGKFNPICKASGWPPTNEEMTLVKPDTVLHPKDFNHTDHSVAFTTHRIILPDSG 420
Db      361 vnsghfnpicakagwpptneemtivkpdgtvlpqndfnytdrfsvaitvnrviipdsq 420
QY      421 VWCSVNTVAGMVKFPNISVKV:PKPLNAPNVIDTGHNFPAVINISSEPFYC----- 472
Db      421 vwcsvntvagmvekpfnisvkv:peplhapnvldtghnfalinssepfyfgdpikskk 480
QY      473 -----EPKS-----CDKTHCTPCPAPPELLGSP-----S 496
Db      481 lfykpvndqakylvctneiftlnyletrdyele---vglarpgqegglpgvrrfttas 538
QY      497 VFLPFPKPKDITLMISRTIP-EVTCVVVDVSHEDPEVKENWYVDGVVHNAKTKPREQYNS 555
Db      539 iglfpfpkglslpksqalntlwpilfused-----ely---veverisqlttdqgn- 589
QY      556 TYKVVSVLT-VLHQDWLNKGYCKKVSNNKALPAPIEKTISKAKGQ-PRHQVYVY----LP 609
Db      590 -ikvpqnltsvllsnlvpreqyvra-----rvntkagqewseclrawtldslilp 638
QY      610 PSKEEM-----TKNVSLTCLVKGFYPSDIAVEWESNQPENNY 648
Db      639 ppenikisnitdstamvswt-ivdgysissliirykvqgknedqh 683

```



```

Query Match      62.4%; Score 2413.5; ID 21; Length 1122;
Best Local Similarity 68.0%; Pred No 4 20-138;
Matches 480; Conservative 50; Mismatches 95; Indels 81; Gaps 14;

OY 1 MDSLASVLCVSVLLISCTVEGAMDLILINSIPVSAFTSLTCTASGWPHPHPITIGRD 60
Db 1 mdsilagivcvsliiyvgvqamdliilnsiplvsdaetstctasgwhphtigrd 60

OY 61 FEALMNHQDPLEVIGDVTREWAKKVVWKEKASKINRAVVEGVPVPEAIPITMKMPQ 120
Db 61 fealmnqhdpvleqvgtvrewakkkvwwkrekaskingayfvegvipvpeaipitmkmrq 120

OY 121 QASFLPATLITMIVUKGDNVNIISKVLIKPEDAVIYKNGSFTHSVPRHEVPDILKVLHP 180
Db 121 qasflpatlmtvdrgdnvnislkvlkpedavlyknsgfthsvprhevpdilevhlph 180

OY 241 ICPPGEMGRTCEKAELEHTFRTCKERCSGQEGCKSYVFLPDYPYGCSCATGKGLQCN 300
Db 241 icppgfmgrtceakaelhtftrtckercsgqegcksyvflpdygcsccatgkglqcn 300

OY 301 ACHPGFYCPDCKLRCSNNGEMCDRFGCLCSFGWGLQCRKGIIPRMTPKIVDLPDHL 360
Db 301 achpgfycpdccklrscnngemcdrfgcclcsfgwglqcrkgiiprmtpkivdlpdlhl 360

OY 361 VNSKFNPICKASGWPPIPTNEEMTLVKKPLNAPNVLDTHGNEFAVINISSEPYFG 420
Db 361 vnskfnpicckasgwpptneemtlvkkplnapnvldthgnefavinisseyfg 420

OY 421 VWVGSVNTVAGMVEKPNISVKVLPKPLNAPNVLDTHGNEFAVINISSEPYFG 480
Db 421 vwvgsvntvagmvekpnisvkvlpkplnapnvldthgnefavinisseyfg 480

OY 473 -----EPKS-----CDKTHTCPPCPAPPELLGGP-----S 496
Db 473 -----EPKS-----CDKTHTCPPCPAPPELLGGP-----S 496

OY 481 ILYKPVQAWKYIEVTEITLYLEPTDYEIC--VQLARPGEGEQHPQVRRITLAC 538
Db 481 ilypvqawkyievtelilyleptdyeic--vqlarpggegqhpqvrrittac 538

OY 497 VFLPPPKKDTLMTLSRTP-EVTCVVVDVSHEDPEVKFNMYVDCGVEYHNAKTPRREQYNS 555
Db 497 vflppppkdtlmtlsrtp-evtcvvvdvshedpevkfnmyvdcgveyhnaaktprreqyns 555

OY 539 IGLPPRGLSLPKSQALNLTWQPIFNSED-----EFY----VEVERTSIGTSDGQN- 589
Db 539 iglppprglslpksqalnlwtqpfinsed-----efy----vevertsigtsgdqn- 589

OY 556 TYPVSVLIT-VLHQDWLWLNCKEYKCVSNKALPAPIERTISKAKGQ-PRPEQVIT-----LP 609
Db 556 typvsvlit-vlhwqdwlnwlnckeykcvsnkalpapiertiskakgq-prpeqvitt-----lp 609

OY 590 -IKYVGNLTSVLSNLYPQDYTVRA-----RVNTKADGEWSCELRATWLSLIP 638
Db 590 -ikvgnltsvlsnlypqdytvra-----rvntkadgewseclratwlsdilp 638

OY 610 PSREFM-----PKNOVSLTCIVKCHYPSPDIADVWESNCOPIENNY 648
Db 610 psrefm-----pknovsltcivkchypspdiadvwesncopienny 648

OY 639 POPENIKSLNITDSTANVSWL-IDVQYSISSIIIRYKVQAKNEDQH 683
Db 639 popenikslnitdstanvswl-idvqysissiiirykvqaknedqh 683

RESULT 11
ID AAR48627
XX AAR48627 standard; Protein; 1122 AA.
AC AAR48627;
XX
XX
XX 31-JUL-1994 (first entry)
XX Protein tyrosine-kinase.
XX Protein-tyrosine-kinase; tek; transcriptional regulatory element;
XX endothelium; gene therapy; vascular disease.
XX
XX OS Mus musculus.
XX
XX PN W09404694-A.
XX
XX PD 03-MAR-1994.
XX

```

```

PF 25-AUG-1993; 93WO-CA00352.
XX
XX 25-AUG-1992; 92WS-0034493.
XX
XX (MOUN ) MOUNT SINAI HOSPITAL CORP.
XX
XX Breitman MJ., Dumont D., Gradwohl GC;
XX
XX WPI: 1994-083208/10.
XX
XX N-PSDB; AAQ56697.
XX
XX Isolated transcriptional regulatory element - for directing
XX expression of gene specifically in cells of endothelial lineage
XX
XX Disclosure; Page 42-46; 74pp; English.
XX
XX cDNA of sequence AAQ56697, derived from mouse strain CD-1 embryo
XX heart, encodes a protein tyrosine-kinase (tek) of sequence AAR48627
XX that is expressed during mouse cardiogenesis. An additional cDNA
XX sequence, given in AAQ56698 and encoding protein AAR48628, was isolated
XX from a phage lambda-gt10 mouse embryo cDNA library. DNA of sequence
XX AAQ56699 was isolated from a mouse genomic phage library using tek
XX cDNA as probe, and included the transcriptional regulatory element
XX of tek comprising the initiation codon and untranslated sequences.
XX The element may be used in gene therapy to introduce foreign genes
XX into endothelial cells to correct/prevent vascular disease.
XX
XX Sequence 1122 AA;
XX
Query Match      62.4%; Score 2413.5; ID 15; Length 1122;
Best Local Similarity 67.7%; Pred No. 2.9e-137;
Matches 478; Conservative 50; Mismatches 97; Indels 81; Gaps 14;

OY 1 MDSLASVLCVSVLLISGIVEGAMDLILINSIPVSAFTSLTCTASGWPHPHPITIGRD 60
Db 1 mdsilagivcvsliiyvgvqamdliilnsiplvsdaetstctasgwhphtigrd 60

OY 61 FEALMNHQDPLEVIGDVTREWAKKVVWKEKASKINRAVVEGVPVPEAIPITMKMPQ 120
Db 61 fealmnqhdpvleqvgtvrewakkkvwwkrekaskingayfvegvipvpeaipitmkmrq 120

OY 121 QASFLPATLITMIVUKGDNVNIISKVLIKPEDAVIYKNGSFTHSVPRHEVPDILKVLHP 180
Db 121 qasflpatlmtvdrgdnvnislkvlkpedavlyknsgfthsvprhevpdilevhlph 180

OY 181 ACPYDAGVYSARYIYGNLFTSAFTLIVPKPEAKWPEINHLTA'MNNVYVHEDTGE' 240
Db 181 acpydagvysaryiygnlftsaftlrvtrceagkwgpcsrpctckngvchcdtge' 240

OY 241 ICPPGEMGRTCEKAELEHTFRTCKERCSGQEGCKSYVFLPDYPYGCSCATGKGLQCN 300
Db 241 icppgfmgrtceakaelhtftrtckercsgqegcksyvflpdygcsccatgkglqcn 300

OY 301 ACHPGFYCPDCKLRCSNNGEMCDRFGCLCSFGWGLQCRKGIIPRMTPKIVDLPDHL 360
Db 301 achpgfycpdccklrscnngemcdrfgcclcsfgwglqcrkgiiprmtpkivdlpdlhl 360

OY 361 VNSKFNPICKASGWPPIPTNEEMTLVKKPLNAPNVLDTHGNEFAVINISSEPYFG 420
Db 361 vnskfnpicckasgwpptneemtlvkkplnapnvldthgnefavinisseyfg 420

OY 421 VWVGSVNTVAGMVEKPNISVKVLPKPLNAPNVLDTHGNEFAVINISSEPYFG 480
Db 421 vwvgsvntvagmvekpnisvkvlpkplnapnvldthgnefavinisseyfg 480

OY 473 -----EPKS-----CDKTHTCPPCPAPPELLGGP-----S 496
Db 473 -----EPKS-----CDKTHTCPPCPAPPELLGGP-----S 496

OY 481 ILYKPVQAWKYIEVTEITLYLEPTDYEIC--VQLARPGEGEQHPQVRRITLAC 538
Db 481 ilypvqawkyievtelilyleptdyeic--vqlarpggegqhpqvrrittac 538

OY 497 VFLPPPKKDTLMTLSRTP-EVTCVVVDVSHEDPEVKFNMYVDCGVEYHNAKTPRREQYNS 555
Db 497 vflppppkdtlmtlsrtp-evtcvvvdvshedpevkfnmyvdcgveyhnaaktprreqyns 555

OY 539 IGLPPRGLSLPKSQALNLTWQPIFNSED-----EFY----VEVERTSIGTSDGQN- 589
Db 539 iglppprglslpksqalnlwtqpfinsed-----efy----vevertsigtsgdqn- 589

```



QY 497 VLFPPKPKDLMISRTIP-EVTCVVVDVSHEDPEVKENWYVDGVEVINAKTKPREQVNS 555  
 DB 538 IGLPPRGISLPKSTALNLTWQPTFNSD-----EFY---VEVERRSIGTSDQGN- 588  
 QY 556 TYRVSVLT-VLIHODWLNKGEYCKVSNKALPAPIEXTISKAKQ-QPREQVYT---LP 609  
 DB 589 -IKVPQNLTVLLNLPREDYVRA-----RVNKAQDEWSEELRAWLISLIP 637  
 QY 610 PSREHM-----TKNOVSIHCLVKGPSPDIAVFWESNGOPENNY 648  
 DB 638 POPENIKSLNLTSTAMVSWL-IVDGYSSIIIRYKVKQKNEQNH 682

RMSJUL 13  
 AAW83160  
 ID AAW83160 standard: Protein, 1101 AA.  
 AC AAW83160;  
 XX  
 DT 11-FEB-1999 (first entry)  
 DE Rat orphan tyrosine kinase receptor Protein Rtk-7 (10c-2).  
 KW Receptor tyrosine kinase, Ror-1; Ror-2, Etk-1, Etk-2, detection;  
 KW neurotrophin activity; TrkB, proto-oncogene, tyrosine kinase receptor;  
 KW binding protein; RORF; NR-3; diagnosis.  
 XX  
 OS Rattus sp.  
 XX US5843749-A.  
 XX 01-DEC-1998.  
 XX 06-JUN-1995; 95US-0469537.  
 XX 17-MAR-1995; 95US-0469247.  
 PR 26-JUL-1991; 91US-0736559.  
 PR 28-OCT-1993; 93US-0144992.  
 PR 06-JUN-1995; 95US-0469537.  
 XX  
 PA (REG-) REGENERON PHARM INC.  
 XX  
 PI Maisompierre PC, Masiakowski P, Yancopoulos GD;  
 XX WPI; 1999-044584/04.  
 DR  
 XX DNA encoding receptor tyrosine kinase proteins - and corresponding  
 PT proteins  
 PT  
 PS Example, Fig 20; 194pp; English.  
 XX  
 CC The present invention describes nucleic acid molecules for Ror-1,  
 CC Ror-2, Etk-1 and Etk-2. Also described are the corresponding proteins  
 CC for Ror-1, Ror-2, Etk-1, and Etk-2. The proteins are orphan receptor  
 CC tyrosine kinases. The present sequence represents rat orphan tyrosine  
 CC kinase receptor Rtk-7 (Tic 2) from the present invention.  
 XX  
 SQ Sequence 1101 AA;

Query Match 61.8%, Score 2388.5, DB 20, Length 1101;  
 Best Local Similarity 69.2%, Pred. No. 9, 20-136;  
 Matches 473; Conservative 40, Mismatches 96; Indels 75; Gaps 13;

QY 20 VEGAMDILLINSLPLVSDAETSLTICIASGWREPEITIGKUFKALMNHQJDPLEVYDVT 79  
 DB 1 VEGAMDILLINSLPLVSDAETSLTICIASGWREPEITIGKUFKALMNHQJDPLEVYDVT 60  
 QY 80 REWAKKVVWKKPKASKINCAVFCGRVHGEAIRTKMKRQOASFLPATLTMTVUKGUNV 139  
 DB 61 REWAKKVVWKKPKASKINCAVFCGRVHGEAIRTKMKRQOASFLPATLTMTVUKGUNV 120

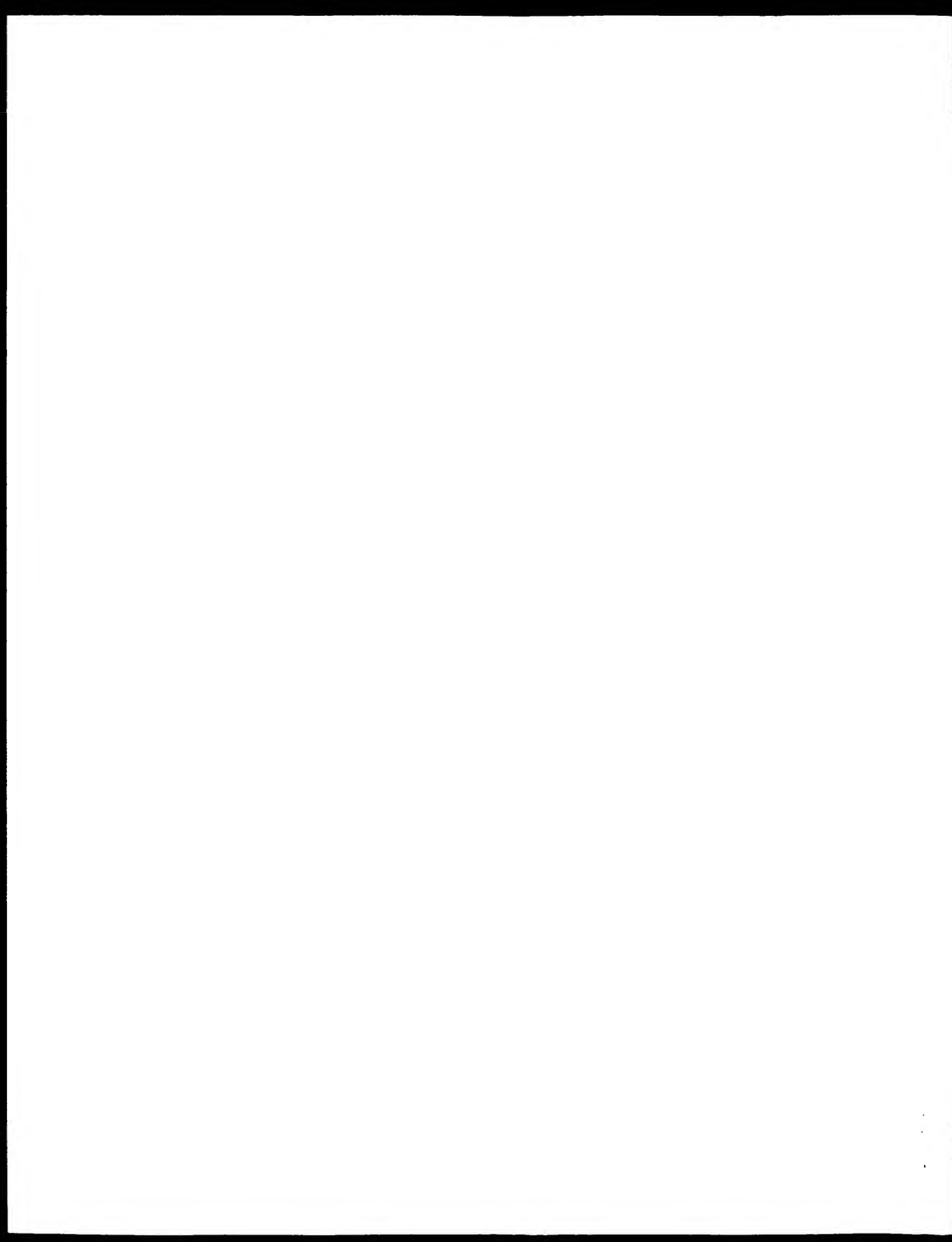
QY 140 NISPKKVLKEEDAVIYKNGSFTHSPYRHEVPEIDLEVHLPHPAQPDAGVYSAPYIGGNLF 199  
 DB 121 NISFKKVLKEEDAVIYKNGSFTHSPYRHEVPEIDLEVHLPHPAQPDAGVYSAPYIGGNLF 180  
 QY 200 TSAPTPLIVPPPCAKWGPENHILCTACMNGVCHEDTGKICPPCFMGRFCHKACHLHT 259  
 DB 181 TSAITLIVRRECAQKWPQDNCNPCTCKNGSVCHEDTGKICPPCFMGRFCHKACHLHT 240  
 QY 260 FGRTCKRCSQKQPKCKSVYVCLIPYVCSACAGWKILWENKACHPGVGPDPCKLPSCNN 319  
 DB 241 FRTCKRCSQKQPKCKSVYVCLIPYVCSACAGWKILWENKACHPGVGPDPCKLPSCNN 300  
 QY 320 GEMTDFPQRTASPRQWQRTQEPESGIPRMTPKIVLLEPHLEVNNSGFENPICKASGWPDP 379  
 DB 301 GEMTDFPQRTASPRQWQRTQEPESGIPRMTPKIVLLEPHLEVNNSGFENPICKASGWPDP 360  
 QY 380 NEEMILVKPIGIVLHKRDNHIDHFSVAFTLHKILPPQSGVWVCSVNTVACMVERKPN 439  
 DB 361 SEEMILVKPIGIVLHKRDNHIDHFSVAFTLHKILPPQSGVWVCSVNTVACMVERKPN 420  
 QY 440 SVKVLKPIGIVLHKRDNHIDHFSVAFTLHKILPPQSGVWVCSVNTVACMVERKPN 484  
 DB 421 SVKVLKPIGIVLHKRDNHIDHFSVAFTLHKILPPQSGVWVCSVNTVACMVERKPN 480  
 QY 485 -----PCAPPELLGSP-----SVFLEPPPKPKDTLMISKIP-EV 515  
 DB 481 VLIYLEPTDYELVGLVPGGEGGHPGVITFTASILPPLPGISLPKSGTALN 540  
 QY 517 TCVVVDVSHEDPEVKENWYVDGVEVINAKTKPREQVNSTYRVSVLT-VLIHODWLNKGE 575  
 DB 541 TWQPTFNSD-----EFY---VEVERRSIGTSDQGN- 590  
 QY 576 YKCKSNKALPAPIEXTISKAKQ-QPREQVYT---LPPSRDEM-TKNOVSIHCLVK-- 627  
 DB 591 YSVRA RVTKAGGQSEELRAWLISLIP 640  
 QY 628 ---GFYPSDIAVFWESNGOPENNY 648  
 DB 641 IVDGYSSIIIRYKVKQKNEQNH 664

RESULT 14  
 AAY68949  
 ID AAY68949 standard; Protein, 764 AA  
 XX  
 AC AAY68949;  
 XX  
 DT 30-MAY-2000 (first entry)  
 DE Fusion protein of murine delta-related protein and human IgG Fc.  
 XX  
 KW Cell development cycle; Delta family, membrane surface-bound ligand;  
 KW endothelial cell biology; gene therapy; subcutaneous infarct;  
 KW cerebral autosomal dominant arteriopathy; leukoencephalopathy;  
 XX ischemic stroke; chimera.  
 OS Chimeric - Mus sp.  
 OS Chimeric - Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT Protein 1..529  
 FT /note= "extracellular region of the murine  
 FT Protein 533..764  
 FT /note= "human IgG Fc portion"  
 XX W0200006725-A2.  
 XX 10-FEB-2000.  
 XX 12-JUL-1999; 99WO-US15710.  
 XX 27-JUL-1998; 98US-0123168.









GenCore version 4.5  
Copyright (c) 1993 - 2000 Genengen Inc.

OM protein - protein search, using sw model

Run on: September 23, 2002, 22:03:19 : Search time 25.42 seconds  
(without alignments)  
676.460 Million cell updates/sec

Title: US-09-733-764-2  
Perfect score: 3868  
Sequence: 1 MSLASLVLCGVSLLSGTV. MIEALHNIYTKSLSLSPK 704

Scoring table: HLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA: \*  
1: /cgn2\_6/prodata/1/aa/5A\_COMB.pep.\*  
2: /cgn2\_6/prodata/1/aa/5B\_COMB.pep.\*  
3: /cgn2\_6/prodata/1/aa/6A\_COMB.pep.\*  
4: /cgn2\_6/prodata/1/aa/6B\_COMB.pep.\*  
5: /cgn2\_6/prodata/1/aa/PC13S\_COMB.pep.\*  
6: /cgn2\_6/prodata/1/aa/bankilles1.pep.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2623	67.8	1124	1	US-08-323-474-2
2	2623	67.8	1124	2	US-09-296-005-16
3	2427.5	62.8	1138	1	US-07-934-393B-2
4	2427.5	62.8	1138	1	US-08-378-089A-2
5	2427.5	62.8	1138	2	US-08-838-957A-2
6	2427.5	62.8	1122	1	US-08-278-089A-6
7	2427.5	62.8	1122	2	US-08-838-957A-6
8	2488.5	61.8	1131	2	US-08-460-537A-6
9	1312	33.9	911	2	US-08-484-438-10
10	1300	33.6	518	1	US-08-385-229-4
11	1298.5	33.6	486	1	US-08-243-010-1
12	1279.5	33.1	396	2	US-08-784-512-3
13	1279.5	33.1	396	4	US-09-176-228-3
14	1276.5	33.0	438	1	US-08-047-827-11
15	1276.5	33.0	438	1	US-08-404-574-11
16	1274	32.9	424	5	PCT-US95-03856-12
17	1274	32.9	424	5	PCT-US95-03856-14
18	1268	32.8	451	2	US-08-887-352B-14
19	1268	32.8	451	2	US-08-287-352B-14
20	1268	32.8	451	2	US-08-887-352B-18
21	1268	32.8	451	3	US-08-466-151-65
22	1268	32.8	451	4	US-09-189-207C-14
23	1268	32.8	451	4	US-09-189-207C-16
24	1268	32.8	451	4	US-09-109-207C-18
25	1268	32.8	451	4	US-09-282-505-2
26	1268	32.8	451	4	US-09-864-266-2
27	1268	32.8	451	4	US-09-296-005-14

Sequence 16, Appl  
Sequence 18, Appl  
Sequence 71, Appl  
Sequence 71, Appl  
Sequence 8, Appl  
Sequence 8, Appl  
Sequence 22, Appl  
Sequence 22, Appl  
Sequence 22, Appl  
Sequence 4, Appl  
Sequence 18, Appl  
Sequence 18, Appl  
Sequence 9, Appl  
Sequence 11, Appl  
Sequence 2, Appl  
Sequence 2, Appl  
Sequence 15, Appl

ALIGNMENTS

RESULT 1  
US-08-323-474-2  
: Sequence 2, Application US/08323474  
: Patent No. 5447860  
: GENERAL INFORMATION:  
: APPLICANT: Ziegler, Steven F.  
: TITLE OF INVENTION: NOVEL TYROSINE KINASE  
: NUMBER OF SEQUENCES: 8  
: CORRESPONDENCE ADDRESS:  
: ADDRESS: Immunex Corporation  
: STREET: 51 University Street  
: CITY: Seattle  
: STATE: Washington  
: COUNTRY: US  
: ZIP: 98101  
: COMPUTER READABLE FORM:  
: MEDIUM TYPE: Floppy disk  
: COMPUTER: IBM PC compatible  
: OPERATING SYSTEM: PC-DOS/MS-DOS  
: SOFTWARE: PatentIn Release #1.0, Version #1.25  
: CURRENT APPLICATION DATA:  
: APPLICATION NUMBER: US/08/323,474  
: FILING DATE:  
: CLASSIFICATION: 435  
: PRIOR APPLICATION DATA:  
: APPLICATION NUMBER: US 07/905,600  
: FILING DATE: 26 JUN 1992  
: ATTORNEY/AGENT INFORMATION:  
: NAME: Seese, Kathryn A.  
: REGISTRATION NUMBER: 32,172  
: REFERENCE/EXHIBIT NUMBER: 2509  
: TELECOMMUNICATION INFORMATION:  
: TELEPHONE: (206) 587-0430  
: TELEFAX: (206) 233-0644  
: TELEX: 756822  
: INFORMATION FOR SEQ ID NO. 2:  
: SEQUENCE CHARACTERISTICS:  
: LENGTH: 1124 amino acids  
: TYPE: amino acid  
: TOPOLOGY: linear  
: MOLECULE TYPE: protein  
US 08-323-474-2

Query Match: 67.8%, Score 2623, Ls 1, Length 1124;  
Best local Similarity: 77.18; Pred No. 2, 80-199;  
Matches 501, Conservative 23, Mismatches 44; Indels 82; Gaps 7;  
: MSLASLVLCGVSLLSGTVSAMLLIINSLFLVSLALSLICLASQWRGRIITGGD 60  
: |||||||||||||||||||||||||||||||||||||||||||||||||||||

```

Db 1 MDLASLVLSVSLLSSTVEGAMLLILNSIPVSDAETSLTCTASGWRPHIEPTIGRD 60
QY 61 FPAIMNQHODPLEVTDVIRWAKVVMKPKASKINCAVEFCGVRGFAIPITMKMPQ 120
Db 61 FPAIMNQHODPLEVTDVIRWAKVVMKPKASKINCAVEFCGVRGFAIPITMKMPQ 120
QY 121 QASFLPATLTMTVDKGDVNIISFKKVLKEEDAVIYKNGSFIHSVPRIEVPDILEVILPH 180
Db 121 QASFLPATLTMTVDKGDVNIISFKKVLKEEDAVIYKNGSFIHSVPRIEVPDILEVILPH 180
QY 181 AQPDQAGVVSARYIGCNLFSAFTPLIVRPGCAKWKPE'NHLTATACMNNVCHIEDTGC 240
Db 181 AQPDQAGVVSARYIGCNLFSAFTPLIVRPGCAKWKPE'NHLTATACMNNVCHIEDTGC 240
QY 241 ICPPGFMGRTCKACELHFGTCKERCSCQGGCKSVYFCPLDPYGCSCATGKGLQCNF 300
Db 241 ICPPGFMGRTCKACELHFGTCKERCSCQGGCKSVYFCPLDPYGCSCATGKGLQCNF 300
QY 301 ACHPEFYGPCKLRSCNNGEMCDREFGGLSPQWGLQ'EREGIPRMTKIVDLPDIE 360
Db 301 ACHPEFYGPCKLRSCNNGEMCDREFGGLSPQWGLQ'EREGIPRMTKIVDLPDIE 360
QY 361 VNSGKFNPICKASGWPPTNEMLVKPDGTVLHPKDNHTDHSVAIFTIHRILPPDSG 420
Db 361 VNSGKFNPICKASGWPPTNEMLVKPDGTVLHPKDNHTDHSVAIFTIHRILPPDSG 420
QY 421 VWCSTNTVAGVVEKPPNISVKVLKPLNAPNVIDTCHNFAVINISSEPYGEPKSCDKT 480
Db 421 VWCSTNTVAGVVEKPPNISVKVLKPLNAPNVIDTCHNFAVINISSEPYGEPKSCDKT 480
QY 481 HTCPCPAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNHYV 536
Db 481 HTCPCPAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNHYV 536
QY 537 --DRVEVINAKTKPREQVNSTYRVVSVLTVLHQWLNKREYKVKSNKALPAPIETIS 594
Db 537 --DRVEVINAKTKPREQVNSTYRVVSVLTVLHQWLNKREYKVKSNKALPAPIETIS 594
QY 595 KAKGQRPPEVYTIIDPSR--PFMKNQVSLCTIVKGYPS---DIAVWE 639
Db 595 KAKGQRPPEVYTIIDPSR--PFMKNQVSLCTIVKGYPS---DIAVWE 639
QY 599 ASIG-----LPPRGLNLLPKSQTLLNTWQPIPPSDEDDFYVEVE 579
Db 599 ASIG-----LPPRGLNLLPKSQTLLNTWQPIPPSDEDDFYVEVE 579

RESULT 3
PCT-US93-06093-2
: Sequence 2, Application US/07934393H
: Patent No. 5466596
: GENERAL INFORMATION:
: APPLICANT: BREITMAN, MARTIN L.
: APPLICANT: DUMONT, DANIEL
: APPLICANT: GRAUHOHL, GERARD G.
: TITLE OF INVENTION: TISSUE SPECIFIC TRANSCRIPTIONAL
: REGULATORY ELEMENT
: NUMBER OF SEQUENCES: 5

```

```

: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (206) 587-0430
: TELEFAX: (206) 233-0644
: TELEX: 756822
: INFORMATION FOR SEQ ID NO: 2:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1124 amino acids
: TYPE: AMINO ACID
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: PCT-US93-06093-2

Query Match 67.8% Score 2623; DB 5; Length 1124;
Best Local Similarity 77.1%; Pred. No. 2,80-199;
Matches 501; Conservative 23; Mismatches 44; Indels 82; Gaps 7;

QY 1 MDLASLVLSVSLLSSTVEGAMLLILNSIPVSDAETSLTCTASGWRPHIEPTIGRD 60
Db 1 MDLASLVLSVSLLSSTVEGAMLLILNSIPVSDAETSLTCTASGWRPHIEPTIGRD 60
QY 61 FPAIMNQHODPLEVTDVIRWAKVVMKPKASKINCAVEFCGVRGFAIPITMKMPQ 120
Db 61 FPAIMNQHODPLEVTDVIRWAKVVMKPKASKINCAVEFCGVRGFAIPITMKMPQ 120
QY 121 QASFLPATLTMTVDKGDVNIISFKKVLKEEDAVIYKNGSFIHSVPRIEVPDILEVILPH 180
Db 121 QASFLPATLTMTVDKGDVNIISFKKVLKEEDAVIYKNGSFIHSVPRIEVPDILEVILPH 180
QY 181 AQPDQAGVVSARYIGCNLFSAFTPLIVRPGCAKWKPE'NHLTATACMNNVCHIEDTGC 240
Db 181 AQPDQAGVVSARYIGCNLFSAFTPLIVRPGCAKWKPE'NHLTATACMNNVCHIEDTGC 240
QY 241 ICPPGFMGRTCKACELHFGTCKERCSCQGGCKSVYFCPLDPYGCSCATGKGLQCNF 300
Db 241 ICPPGFMGRTCKACELHFGTCKERCSCQGGCKSVYFCPLDPYGCSCATGKGLQCNF 300
QY 301 ACHPEFYGPCKLRSCNNGEMCDREFGGLSPQWGLQ'EREGIPRMTKIVDLPDIE 360
Db 301 ACHPEFYGPCKLRSCNNGEMCDREFGGLSPQWGLQ'EREGIPRMTKIVDLPDIE 360
QY 361 VNSGKFNPICKASGWPPTNEMLVKPDGTVLHPKDNHTDHSVAIFTIHRILPPDSG 420
Db 361 VNSGKFNPICKASGWPPTNEMLVKPDGTVLHPKDNHTDHSVAIFTIHRILPPDSG 420
QY 421 VWCSTNTVAGVVEKPPNISVKVLKPLNAPNVIDTCHNFAVINISSEPYGEPKSCDKT 480
Db 421 VWCSTNTVAGVVEKPPNISVKVLKPLNAPNVIDTCHNFAVINISSEPYGEPKSCDKT 480
QY 481 HTCPCPAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNHYV 536
Db 481 HTCPCPAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNHYV 536
QY 537 --DRVEVINAKTKPREQVNSTYRVVSVLTVLHQWLNKREYKVKSNKALPAPIETIS 594
Db 537 --DRVEVINAKTKPREQVNSTYRVVSVLTVLHQWLNKREYKVKSNKALPAPIETIS 594
QY 595 KAKGQRPPEVYTIIDPSR--PFMKNQVSLCTIVKGYPS---DIAVWE 639
Db 595 KAKGQRPPEVYTIIDPSR--PFMKNQVSLCTIVKGYPS---DIAVWE 639
QY 599 ASIG-----LPPRGLNLLPKSQTLLNTWQPIPPSDEDDFYVEVE 579
Db 599 ASIG-----LPPRGLNLLPKSQTLLNTWQPIPPSDEDDFYVEVE 579

RESULT 3
US-07-934-493H-2
: Sequence 2, Application US/07934393H
: Patent No. 5466596
: GENERAL INFORMATION:
: APPLICANT: BREITMAN, MARTIN L.
: APPLICANT: DUMONT, DANIEL
: APPLICANT: GRAUHOHL, GERARD G.
: TITLE OF INVENTION: TISSUE SPECIFIC TRANSCRIPTIONAL
: REGULATORY ELEMENT
: NUMBER OF SEQUENCES: 5

```

? CORRESPONDENCE ADDRESS:  
 ? ADDRESSEE: BERESKIN & PARR  
 ? STREET: 40 King Street West  
 ? CITY: Toronto  
 ? STATE: Ontario  
 ? COUNTRY: Canada  
 ? ZIP: M5H 3Y2  
 ? COMPUTER READABLE FORM:  
 ? MEDIUM TYPE: Floppy disk  
 ? COMPUTER: IBM PC compatible  
 ? OPERATING SYSTEM: PC-DOS/MS-DOS  
 ? SOFTWARE: PatentIn Release #1.0, Version #1.25  
 ? CURRENT APPLICATION DATA:  
 ? APPLICATION NUMBER: US/07/934,393R  
 ? FILING DATE: 25-AUG-1992  
 ? CLASSIFICATION: 435  
 ? ATTORNEY/AGENT INFORMATION:  
 ? NAME: Kurdydyk, Linda M.  
 ? REGISTRATION NUMBER: 34,971  
 ? REFERENCE/DOCKET NUMBER: 3153-64  
 ? TELECOMMUNICATION INFORMATION:  
 ? TELEPHONE: (416) 354-7311  
 ? TELEFAX: (416) 361-1398  
 ? INFORMATION FOR SEQ ID NO: 2:  
 ? SEQUENCE CHARACTERISTICS:  
 ? LENGTH: 1118 amino acids  
 ? TYPE: amino acid  
 ? TOPOLOGY: linear  
 ? MOLECULE TYPE: protein  
 ? US-07-934-393B-2

Query Match 62.8%; Score 2427.5; DB 1; Length 1118;  
 Best Local Similarity 68.0%; Pred. No. 8 3e-184;  
 Matches 480; Conservative 50; Mismatches 95; Indels 81; Gaps 14,  
 QY 1 MDSLASLVLCGVSLLSSTVFAGMILLINSLPLVSADTSLT-IASGWPPEPTIGRD 60  
 DB 1 MDSLAGVLCVSLLVGVVGVGAMILLINSLPLVSADTSLT-IASGWPPEPTIGRD 60  
 QY 61 FEALMNOHQDPLEVQIVTFRWAKKVVWKKREKASK INCAYFCGRVKEA IRI RTKMRQ 120  
 DB 61 FEALMNOHQDPLEVQIVTFRWAKKVVWKKREKASK INCAYFCGRVKEA IRI RTKMRQ 120  
 QY 121 QASFLPATLITMTVDKGNVNI SFKKVLIKEEDAVIYKNGSFHSVPREHVPDILEVHLPH 180  
 DB 121 QASFLPATLITMTVDKGNVNI SFKKVLIKEEDAVIYKNGSFHSVPREHVPDILEVHLPH 180  
 QY 181 AOPQDAVYSARYTGSNLTSAFTPLIVRCEAQKWPCECNHLC-TA-MNNGVCHEDTGE 240  
 DB 181 AOPQDAVYSARYTGSNLTSAFTPLIVRCEAQKWPCECNHLC-TA-MNNGVCHEDTGE 240  
 QY 241 ICPGFMGRICFKACFTHIFGPIKFWCSQDPKSKSVVPCIPHPYGSATSWKGLQCNE 300  
 DB 241 ICPGFMGRICFKACFTHIFGPIKFWCSQDPKSKSVVPCIPHPYGSATSWKGLQCNE 300  
 QY 301 ACHPCYGPICFKICSCNNGHMCDFGQICSPGQGLQCCEPCIPPMTHKIVLPHPHIE 360  
 DB 301 ACHPCYGPICFKICSCNNGHMCDFGQICSPGQGLQCCEPCIPPMTHKIVLPHPHIE 360  
 QY 361 VNSGKNPICKASGWLPTNEMLTIVKPDGTVLHPKDFNHTDHFSAVFIITHRLPPDSG 420  
 DB 361 VNSGKNPICKASGWLPTNEMLTIVKPDGTVLHPKDFNHTDHFSAVFIITHRLPPDSG 420  
 QY 421 VVVCSTNTVAGMVEKPNISVKVLPKPLNAPNVIDTGHNAVINISSEPYFG----- 472  
 DB 421 VVVCSTNTVAGMVEKPNISVKVLPKPLNAPNVIDTGHNAVINISSEPYFG----- 472  
 QY 473 -----EPKS-----CDKTHTCPPCPAPELLGCP-----S 496  
 DB 481 LFYKPVNQAWKY IEVTHNIFLTVNLEPRTDYELC--VQLARPGGCGCHCPVRRFTTAC 538  
 QY 497 VLFPPPKDGLMISRTF-EVTCVGVVDVSHEDPEVKFNWVVDGVVEVHNIAKTKPREQYNS 555

DB 539 IGLPPPKGLSLLPKSUALNLTWQPIFTNSD-----EPY-----VEVERSLQTISDQON- 589  
 QY 556 TYRVVSVLT-VLRDWMINGCKEYKCKVSNKALPAP IHTKISKAKGO-PRHPQVYI-----LP 609  
 DB 590 -TRVPRNLTSLVLSNLTVPPEQTVFA-----PVTAKAGWSEELRAWLTLSLILP 638  
 QY 610 PRPEEM-----TRNQVSLTCLVKGFYPSDIAVWESNGQPENNY 648  
 DB 639 PPHENIKISNITDSTAMVSWT-IVVQYSISSIIRKVKQCKKNFVGH 683

## RESULT 4

US-08-278-089A-2  
 ? Sequence 2, Application: US/98278089A  
 ? Patent No. 5681714  
 ? GENERAL INFORMATION:  
 ? APPLICANT: Breiman, Martin L.  
 ? APPLICANT: Rossant, Janet  
 ? APPLICANT: Dumont, Daniel J.  
 ? APPLICANT: Yamaguchi, Terry P.  
 ? TITLE OF INVENTION: NO. 5681714e1 Receptor Tyrosine Kinase  
 ? NUMBER OF SEQUENCES: 33  
 ? CORRESPONDENCE ADDRESS:  
 ? ADDRESSEE: Bereskin & Parr  
 ? STREET: 40 King Street West  
 ? CITY: Toronto  
 ? STATE: Ontario  
 ? COUNTRY: Canada  
 ? ZIP: M5H 3Y2  
 ? COMPUTER READABLE FORM:  
 ? MEDIUM TYPE: Floppy disk  
 ? COMPUTER: IBM PC compatible  
 ? OPERATING SYSTEM: PC-DOS/MS-DOS  
 ? SOFTWARE: PatentIn Release #1.0, Version #1.30  
 ? CURRENT APPLICATION DATA:  
 ? APPLICATION NUMBER: US/98/278,089A  
 ? FILING DATE: 20-JUL-1994  
 ? CLASSIFICATION: 530  
 ? ATTORNEY/AGENT INFORMATION:  
 ? NAME: Kurdydyk, Linda M.  
 ? REGISTRATION NUMBER: 34,971  
 ? REFERENCE/DOCKET NUMBER: 3153-111  
 ? TELECOMMUNICATION INFORMATION:  
 ? TELEPHONE: (416) 364-7311  
 ? TELEFAX: (416) 361-1398  
 ? INFORMATION FOR SEQ ID NO: 2:  
 ? SEQUENCE CHARACTERISTICS:  
 ? LENGTH: 1118 amino acids  
 ? TYPE: amino acid  
 ? TOPOLOGY: linear  
 ? MOLECULE TYPE: protein  
 ? US-08-278-089A-2

Query Match 62.8%; Score 2427.5; DB 1; Length 1118;  
 Best Local Similarity 68.0%; Pred. No. 8 3e-184;  
 Matches 480; Conservative 50; Mismatches 95; Indels 81; Gaps 14-  
 QY 1 MDSLASLVLCGVSLLSSTVFAGMILLINSLPLVSADTSLT-IASGWPPEPTIGRD 60  
 DB 1 MDSLAGVLCVSLLVGVVGVGAMILLINSLPLVSADTSLT-IASGWPPEPTIGRD 60  
 QY 61 FEALMNOHQDPLEVQIVTFRWAKKVVWKKREKASK INCAYFCGRVKEA IRI RTKMRQ 120  
 DB 61 FEALMNOHQDPLEVQIVTFRWAKKVVWKKREKASK INCAYFCGRVKEA IRI RTKMRQ 120  
 QY 121 QASFLPATLITMTVDKGNVNI SFKKVLIKEEDAVIYKNGSFHSVPREHVPDILEVHLPH 180  
 DB 121 QASFLPATLITMTVDKGNVNI SFKKVLIKEEDAVIYKNGSFHSVPREHVPDILEVHLPH 180  
 QY 181 AOPQDAVYSARYTGSNLTSAFTPLIVRCEAQKWPCECNHLC-TA-MNNGVCHEDTGE 240  
 DB 181 AOPQDAVYSARYTGSNLTSAFTPLIVRCEAQKWPCECNHLC-TA-MNNGVCHEDTGE 240

```

181 ACPHAGVYSARYIGGNLTSAFTLIVRRCLAKWGHCSPKICTCKNNVCVCHDTGEC 240
QY 241 ICPHAGVYSARYIGGNLTSAFTLIVRRCLAKWGHCSPKICTCKNNVCVCHDTGEC 240
DB 241 ICPHAGVYSARYIGGNLTSAFTLIVRRCLAKWGHCSPKICTCKNNVCVCHDTGEC 240
QY 301 ACPHAGVYSARYIGGNLTSAFTLIVRRCLAKWGHCSPKICTCKNNVCVCHDTGEC 240
DB 301 ACPHAGVYSARYIGGNLTSAFTLIVRRCLAKWGHCSPKICTCKNNVCVCHDTGEC 240
QY 361 VNSGFNPICAKSGWPLPTNEEMTLVKPDGTVLHPKDFNIHDSVAIFTHIRLPPDSG 420
DB 361 VNSGFNPICAKSGWPLPTNEEMTLVKPDGTVLHPKDFNIHDSVAIFTHIRLPPDSG 420
QY 421 VNSGFNPICAKSGWPLPTNEEMTLVKPDGTVLHPKDFNIHDSVAIFTHIRLPPDSG 420
DB 421 VNSGFNPICAKSGWPLPTNEEMTLVKPDGTVLHPKDFNIHDSVAIFTHIRLPPDSG 420
QY 473 VNSGFNPICAKSGWPLPTNEEMTLVKPDGTVLHPKDFNIHDSVAIFTHIRLPPDSG 420
DB 473 VNSGFNPICAKSGWPLPTNEEMTLVKPDGTVLHPKDFNIHDSVAIFTHIRLPPDSG 420
QY 481 LEYKPVNQAWKYIEVTNEIFTLNLEPKTDYELC--VOLAKPGGGGGHGPVRRFTTAC 538
DB 481 LEYKPVNQAWKYIEVTNEIFTLNLEPKTDYELC--VOLAKPGGGGGHGPVRRFTTAC 538
QY 497 VLFPPKPKDLMISRTP-EVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREEQYNS 555
DB 497 VLFPPKPKDLMISRTP-EVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREEQYNS 555
QY 539 IGLPPHGLSLPKSQTALNLTPQIFTNSED-----EFY---VEVERKRSIQTTSDQON- 589
DB 539 IGLPPHGLSLPKSQTALNLTPQIFTNSED-----EFY---VEVERKRSIQTTSDQON- 589
QY 556 TYRVVSVLT-VLHQDLNCKEYKCKVSNKALPAPIETKISKAKGQ-PRPQVYTI-----LP 609
DB 556 TYRVVSVLT-VLHQDLNCKEYKCKVSNKALPAPIETKISKAKGQ-PRPQVYTI-----LP 609
QY 590 -IKVPGNLTSLVLSNLPVREQTVRA-----RVNTKAUGSESEELRAWTLSLILP 638
DB 590 -IKVPGNLTSLVLSNLPVREQTVRA-----RVNTKAUGSESEELRAWTLSLILP 638
QY 610 PSREEM-----TKNQVSLTCLVKGFYPSDIAVEWISNGQPENNY 648
DB 610 PSREEM-----TKNQVSLTCLVKGFYPSDIAVEWISNGQPENNY 648
QY 639 PGPENIKISNLIUSIAMVSWI-IVIGYSISSIIIRYKVGCKNEDQH 683
DB 639 PGPENIKISNLIUSIAMVSWI-IVIGYSISSIIIRYKVGCKNEDQH 683

```

## RESULT 5

US-08-838-957A-2

Sequence 2: Application US/08/838-957A

Patent No. 5998187

GENERAL INFORMATION:

APPLICANT: Breitman, Martin L.

APPLICANT: Rossant, Janet

APPLICANT: Dumont, Daniel J.

APPLICANT: Yamaguchi, Terry P.

TITLE OF INVENTION: No. 5998187el Receptor Tyrosine Kinase

NUMBER OF SEQUENCES: 32

CORRESPONDENCE ADDRESS:

ADDRESS: Hereskin &amp; Parr

STREET: 40 King Street West

CITY: Toronto

STATE: Ontario

COUNTRY: Canada

ZIP: M5H 3Y2

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/838-957A

FILING DATE: 23-APR-1997

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Kurdydyk, Linda M.

REGISTRATION NUMBER: 34,971

REFERENCE/DOCKET NUMBER: 3153-212

TELECOMMUNICATION INFORMATION:

TELEPHONE: (416) 364-7311

TELEFAX: (416) 361-1398

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 1118 amino acids

TYPE: amino acid

TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-838-957A-2

Query Match 62.8%, Score 2427.5, DB 2, length 1118;

Best local Similarity 68.0%, Pred. No. 8,30-184;

Matches 480; Conservative 50; Mismatches 95; Indels 81; Gaps 14;

QY 1 MSLASLVGVSVLLSGVIGVAGMULLLINSPLVSVIAETSLTICIASGWAPHEPITIGRD 60

DB 1 MSLASLVGVSVLLSGVIGVAGMULLLINSPLVSVIAETSLTICIASGWAPHEPITIGRD 60

QY 61 FEALANQHDPLEVTDVTRWAKKVVWKKASK INCAYPCGKVRGEALRIKIMKMO 120

DB 61 FEALANQHDPLEVTDVTRWAKKVVWKKASK INCAYPCGKVRGEALRIKIMKMO 120

QY 121 QASFLPATLITMTVUGDNVNIISFKKVLIKEDAVIYKNGSP IHSVPKREVDHLEVLPH 180

DB 121 QASFLPATLITMTVUGDNVNIISFKKVLIKEDAVIYKNGSP IHSVPKREVDHLEVLPH 180

QY 181 AUPDAGVYSARYIGGNLTSAFTLIVRRCLAKWGHCSPKICTCKNNVCVCHDTGEC 240

DB 181 AUPDAGVYSARYIGGNLTSAFTLIVRRCLAKWGHCSPKICTCKNNVCVCHDTGEC 240

QY 241 ICPHAGVYSARYIGGNLTSAFTLIVRRCLAKWGHCSPKICTCKNNVCVCHDTGEC 240

DB 241 ICPHAGVYSARYIGGNLTSAFTLIVRRCLAKWGHCSPKICTCKNNVCVCHDTGEC 240

QY 301 ACPHAGVYSARYIGGNLTSAFTLIVRRCLAKWGHCSPKICTCKNNVCVCHDTGEC 240

DB 301 ACPHAGVYSARYIGGNLTSAFTLIVRRCLAKWGHCSPKICTCKNNVCVCHDTGEC 240

QY 361 VNSGFNPICAKSGWPLPTNEEMTLVKPDGTVLHPKDFNIHDSVAIFTHIRLPPDSG 420

DB 361 VNSGFNPICAKSGWPLPTNEEMTLVKPDGTVLHPKDFNIHDSVAIFTHIRLPPDSG 420

QY 421 VNSGFNPICAKSGWPLPTNEEMTLVKPDGTVLHPKDFNIHDSVAIFTHIRLPPDSG 420

DB 421 VNSGFNPICAKSGWPLPTNEEMTLVKPDGTVLHPKDFNIHDSVAIFTHIRLPPDSG 420

QY 473 VNSGFNPICAKSGWPLPTNEEMTLVKPDGTVLHPKDFNIHDSVAIFTHIRLPPDSG 420

DB 473 VNSGFNPICAKSGWPLPTNEEMTLVKPDGTVLHPKDFNIHDSVAIFTHIRLPPDSG 420

QY 481 LEYKPVNQAWKYIEVTNEIFTLNLEPKTDYELC--VOLAKPGGGGGHGPVRRFTTAC 538

DB 481 LEYKPVNQAWKYIEVTNEIFTLNLEPKTDYELC--VOLAKPGGGGGHGPVRRFTTAC 538

QY 497 VLFPPKPKDLMISRTP-EVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREEQYNS 555

DB 497 VLFPPKPKDLMISRTP-EVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREEQYNS 555

QY 539 IGLPPHGLSLPKSQTALNLTPQIFTNSED-----EFY---VEVERKRSIQTTSDQON- 589

DB 539 IGLPPHGLSLPKSQTALNLTPQIFTNSED-----EFY---VEVERKRSIQTTSDQON- 589

QY 556 TYRVVSVLT-VLHQDLNCKEYKCKVSNKALPAPIETKISKAKGQ-PRPQVYTI-----LP 609

DB 556 TYRVVSVLT-VLHQDLNCKEYKCKVSNKALPAPIETKISKAKGQ-PRPQVYTI-----LP 609

QY 590 -IKVPGNLTSLVLSNLPVREQTVRA-----RVNTKAUGSESEELRAWTLSLILP 638

DB 590 -IKVPGNLTSLVLSNLPVREQTVRA-----RVNTKAUGSESEELRAWTLSLILP 638

QY 610 PSREEM-----TKNQVSLTCLVKGFYPSDIAVEWISNGQPENNY 648

DB 610 PSREEM-----TKNQVSLTCLVKGFYPSDIAVEWISNGQPENNY 648

QY 639 PGPENIKISNLIUSIAMVSWI-IVIGYSISSIIIRYKVGCKNEDQH 683

DB 639 PGPENIKISNLIUSIAMVSWI-IVIGYSISSIIIRYKVGCKNEDQH 683

## RESULT 6

US-08-278-089A-6

Sequence 6: Application US/08278089A

Patent No. 5681714

GENERAL INFORMATION:

APPLICANT: Breitman, Martin L.

APPLICANT: Rossant, Janet

APPLICANT: Dumont, Daniel J.

APPLICANT: Yamaguchi, Terry P.

TITLE OF INVENTION: No. 5681714el Receptor Tyrosine Kinase

NUMBER OF SEQUENCES: 33

CORRESPONDENCE ADDRESS:

ADDRESS: Bereskin &amp; Parr

STREET: 40 King Street West

CITY: Toronto

STATE: Ontario  
 COUNTRY: Canada  
 ZIP: M5H 3Y2  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/278,089A  
 FILING DATE: 20-Jul-1994  
 CLASSIFICATION: 530  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Kurdydyk, Linda M.  
 REGISTRATION NUMBER: 34,971  
 REFERENCE/DOCKET NUMBER: 3153-111  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (416) 364-7311  
 TELEFAX: (416) 361-1398  
 INFORMATION FOR SEQ ID NO: 6:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 1122 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 US-08-278-089A-6

Query Match 62.8%; Score 2427.5; DB 1; Length 1122;  
 Best Local Similarity 68.0%; Pred. No. 8.4e-184;  
 Matches 480; Conservative 50; Mismatches 95; Indels 81; Gaps 14.

QY 1 MDSLASLVLCGVSLLSSTVERAMCILLINSLPLVSDAETSLTCLASGWPHEPITIGED 60  
 DB 1 MDSLAGLVLCGVSLLYGVVGAMDELLINSLPLVSDAETSLTCLASGWPHEPITIGED 60  
 QY 61 FEALMNHQHQPLEVTDVTRFWAKKVVWKKKASKINGAYFCGVRGCAIRITMKMQ 120  
 DB 61 FEALMNHQHQPLEVTDVTRFWAKKVVWKKKASKINGAYFCGVRGCAIRITMKMQ 120  
 QY 121 QASFLPATLMTVUKGNVNISFKKVLKEEDAVIYKNGSFHSVPRHEVFDLLEVLPH 180  
 DB 121 QASFLPATLMTVUKGNVNISFKKVLKEEDAVIYKNGSFHSVPRHEVFDLLEVLPH 180  
 QY 181 AQODAGVYSARYTGGNLTSAFTPLIVPCEAQKMGPECNHLCIACMNGVCHEDTGE 240  
 DB 181 AQODAGVYSARYTGGNLTSAFTPLIVPCEAQKMGPECNHLCIACMNGVCHEDTGE 240  
 QY 241 ICPPGFMGRICEKACELHTFORICEKESGVBGCKSYVFCPLDPPYGCSCAIGWKGLQNE 300  
 DB 241 ICPPGFMGRICEKACELHTFORICEKESGVBGCKSYVFCPLDPPYGCSCAIGWKGLQNE 300  
 QY 301 ACHPGFYCPCKLPSTNNKEMTDPPGSLCLSPWQHLQEPERIPPMTPKIVLFDHIE 360  
 DB 301 ACPGYYGPDCKLRCHCTNEICDRFGCLCSOQWGLQCEKEGCRPMTPQIEDLPDIE 360  
 QY 361 VNSGKFNPKCKASGWLPTNEEMTLVPDGTVLHPKDFNHTDHSVAIFTHIRLLPDSG 420  
 DB 361 VNSGKFNPKCKASGWLPTNEEMTLVPDGTVLHPKDFNHTDHSVAIFTHIRLLPDSG 420  
 QY 421 VWGCVNIVAGMVKFPFNISVKVLPKPLNAPNVIDTGNFAVINISSEPYFG----- 472  
 DB 421 VWGCVNIVAGMVKFPFNISVKVLPKPLNAPNVIDTGNFAVINISSEPYFG----- 472  
 QY 473 -----EPKS-----CDKTHCTCPCPAPPELLGGP-----S 496  
 DB 481 LFYKPVQAMKYIEVINEIFLANTLEPRTDYELC--VOLARPGGGCHGCPVRRFTTAC 538  
 QY 497 VFLEPPKPKDTLMSRTP-EVTCVVWVDSHFDEPKFNWVDGVVEVINAKTKPPEOYNS 555  
 DB 539 IGLPPPGSLSLPKSQIALNLITWQPIETNSED-----EFY---VEVEKRSLSQTSQUN- 589  
 QY 556 TYRVVSVLIVLHQLWNGREYKCKVSNKALPAPLEKILSKAKGQ-PRKPVVYI---LP 609

DB 590 -IKVPGNITSVLISNIVPFGYIVAA-----FVNIAQGSSEFLPAWILSLP 638  
 QY 610 PSREEM-----TKNOVSLTCLVKGFYPSDIAVEWESNGOPENNY 648  
 DB 639 PQPNTKISNITDSIAMVSWT-IVNGYSISSIIIPYKVGGRNEQH 683  
 RESULT 7  
 US-08-838-957A-6  
 Sequence 5; Application US/08838957A  
 Patent No.: 5998187  
 GENERAL INFORMATION:  
 APPLICANT: Breitman, Martin L.  
 APPLICANT: Kossant, Janet  
 APPLICANT: Dumont, Daniel J.  
 APPLICANT: Yamaguchi, Terry P.  
 TITLE OF INVENTION: No 5998187a1 Receptor Tyrosine Kinase  
 NUMBER OF SEQUENCES: 32  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Bereskin & Parr  
 STREET: 40 King Street West  
 CITY: Toronto  
 STATE: Ontario  
 COUNTRY: Canada  
 ZIP: M5H 3Y2  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent In Release #1 0, Version #1 30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/838-957A  
 FILING DATE: 23-APR-1997  
 CLASSIFICATION: 435  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Kurdydyk, Linda M.  
 REGISTRATION NUMBER: 34,971  
 REFERENCE/DOCKET NUMBER: 3153-212  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (416) 364-7311  
 TELEFAX: (416) 361-1398  
 INFORMATION FOR SEQ ID NO: 6:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 1122 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 US-08-838-957A-6

Query Match 62.8%; Score 2427.5; DB 2; Length 1122;  
 Best Local Similarity 58.0%; Pred. No. 8.4e-184;  
 Matches 480; Conservative 50; Mismatches 95; Indels 81; Gaps 14;

QY 1 MDSLASLVLCGVSLLSGTGVRGAMDELLINSLPLVSDAETSLTCLASGWPHEPITIGED 60  
 DB 1 MDSLAGLVLCGVSLLYGVVGAMDELLINSLPLVSDAETSLTCLASGWPHEPITIGED 60  
 QY 61 FEALMNHQHQPLEVTDVTRFWAKKVVWKKKASKINGAYFCGVRGCAIRITMKMQ 120  
 DB 61 FEALMNHQHQPLEVTDVTRFWAKKVVWKKKASKINGAYFCGVRGCAIRITMKMQ 120  
 QY 121 QASFLPATLMTVUKGNVNISFKKVLKEEDAVIYKNGSFHSVPRHEVFDLLEVLPH 180  
 DB 121 QASFLPATLMTVUKGNVNISFKKVLKEEDAVIYKNGSFHSVPRHEVFDLLEVLPH 180  
 QY 181 AQODAGVYSARYTGGNLTSAFTPLIVPCEAQKMGPECNHLCIACMNGVCHEDTGE 240  
 DB 181 AQODAGVYSARYTGGNLTSAFTPLIVPCEAQKMGPECNHLCIACMNGVCHEDTGE 240  
 QY 241 ICPPGFMGRICEKACELHTFORICEKESGVBGCKSYVFCPLDPPYGCSCAIGWKGLQNE 300  
 DB 241 ICPPGFMGRICEKACELHTFORICEKESGVBGCKSYVFCPLDPPYGCSCAIGWKGLQNE 300

DB 241 ICPHGMCKICCKACAPHEIFGRICCKRSGSGEGCKSVYFCLDPPYGGSCAIGWKGGLQCNF 300  
QY 301 ACBPGEFYRPUCKI RQSCNN:EMWDPFGCTASPCWGLQ:EPGIGPMTPKIVLDLPDHE 360  
DB 301 ACBPGEFYRPUCKI RQSCNN:EMWDPFGCTASPCWGLQ:EPGIGPMTPKIVLDLPDHE 360  
QY 361 VNSGRFNPKICKASGMPDINMILVAKDGLIVLHFKDIPNHDHFSVAIFTHKILPPDPSG 420  
DB 361 VNSGRFNPKICKASGMPDINMILVAKDGLIVLHFKDIPNHDHFSVAIFTHKILPPDPSG 420  
QY 421 VVWCVSVNTVAGWVEPPFNISVKVLKPKLNADPNVDTGHNFAVINISSEPEFG----- 472  
DB 421 VVWCVSVNTVAGWVEPPFNISVKVLKPKLNADPNVDTGHNFAVINISSEPEFG----- 472  
QY 473 -----EPKS CDKTHICPPCAPPELLGGP-----S 496  
DB 481 LFKVPVQWAKYIEVTNEIFTNILEPTDYELG--VOLAPP-EGEGSHPPVPRPFTIA\* 538  
QY 497 VFLPPTPKKDTLMISRTIP-DVYCVVVDVSHEDHPRKFNWYDGVFVHNAKIKPRFGQYNS 555  
DB 539 IGLPPRGLSLIPKQSQTALNTWQPIETNSEO-----EFY---VEVERPRLQTTSDQCN- 589  
QY 556 TVPVVSVLT-VLHGLWLN-KYKPKVSNKALPAPLETKYAKQO-PPREPQVYT-----LP 609  
DB 590 -IKVPCNTISVTSNIVPPEVTVHA-----PVNTKAGPFWSEFPAMVLSLOPP 638  
QY 610 PSREEM-----TKNWSLTLCKGPEYPSDIAVENESNCPENNY 648  
DB 639 PQPENIKISNITDSTAMYSWT IVLGYSTSSIIIRYKVKQKNEQHH 683

## RESULT R

US-08-469-537A-96

: Sequence 96, Application US/08469537A

: Patent No. 5843749

: GENERAL INFORMATION:

: APPLICANT: Maisonnier, et al.

: TITLE OF INVENTION: HKK AND FOR TYROSINE

: TITLE OF INVENTION: KINASES

: NUMBER OF SEQUENCES: 107

: CORRESPONDENCE ADDRESS:

: ADDRESSER: Regeneron Pharmaceuticals, Inc.

: STREET: 777 Old Saw Mill River Road

: CITY: Tarrytown

: STATE: NY

: COUNTRY: U.S.A.

: ZIP: 10591

: COMPUTER READABLE FORM:

: MEDIUM TYPE: Diskette

: COMPUTER: IBM Compatible

: OPERATING SYSTEM: DOS

: SOFTWARE: FastSeq Version 2.0

: CURRENT APPLICATION DATA:

: APPLICATION NUMBER: US/09/469,537A

: FILING DATE: 06-JUN-1995

: CLASSIFICATION: 435

: PRIOR APPLICATION DATA:

: APPLICATION NUMBER: USSN 08/406,247

: FILING DATE: 17-MAR-1995

: APPLICATION NUMBER: USSN 08/144,992

: FILING DATE: 28-OCT-1993

: APPLICATION NUMBER: USSN 07/736,559

: FILING DATE: 26-JUL-1991

: ATTORNEY/AGENT INFORMATION:

: NAME: Kopleff, Ph.D., Gail M

: REGISTRATION NUMBER: 32,143

: REFERENCE/DOCKET NUMBER: RFG 070C

: TELECOMMUNICATION INFORMATION:

: TELEPHONE: 914-345-7400

: TELEFAX: 914-345-7721

: TELEX:

: INFORMATION FOR SEQ ID NO: 96:

: SEQUENCE CHARACTERISTICS:

: LENGTH: 1101 amino acids  
: TYPE: amino acid  
: STRANDEDNESS: unknown  
: TOPOLOGY: unknown  
: MOLECULE TYPE: protein  
US-08-469-537A-96

Query Match 61.8%; Score 2388.5; DB 2; Length 1101;  
Best Local Similarity 69.2%; Pred. No. 1e-180;  
Matches 473; Conservative 40; Mismatches 96; Indels 75; Gaps 14;

QY 20 VEGAMLLILINSLPLVSDAETSLICIASGMPHEPITIGRDFEALMKQHQHRIIVLQDVT 79

DB 1 VEGAMLLILINSLPLVSDAETSLICIASGMPHEPITIGRDFEALMKQHQHRIIVLQDVT 60

QY 80 RPAWKKVVKPKASKINCAVECPVPVGGAIPIKIMKPDQASFEPAIIMVDSQNV 149

DB 61 RPAWKKVVKPKASKINCAVECPVPVGGAIPIKIMKPDQASFEPAIIMVDSQNV 120

QY 140 NLSKPKVLIKEDDAVIYKNGSEIHSVPRHEVPDILEVILPHAGQPDAGVSARYIGNLF 199

DB 121 NLSKPKVLIKEDDAVIYKNGSEIHSVPRHEVPDILEVILPHAGQPDAGVSARYIGNLF 180

QY 200 TSAFTPIIVRPFACQWPPPCNHICTACMNNVCVCHEDIGSCIDPPPCMOPTDIAKELPI 259

DB 181 TSAFTPIIVRPFACQWPPPCNHICTACMNNVCVCHEDIGSCIDPPPCMOPTDIAKELPI 240

QY 260 FQPT-KKEG-SQAPGPKSVVETIPQYKASCAIHWKGLQ-NEN-HPGFGYGPCKLRHCIN 419

DB 241 FQPTKERCSECKSVVFCIPDPYGCSCAIGWRGLQCNACPYGHYGPCKLRHCIN 400

QY 320 REMTDFQNT-SP-SW-SLQ-TEPEELPMTFKIVLEPHIEVNSKFNPKKASGMPLEPT 479

DB 301 REMTDFQNT-SP-SW-SLQ-TEPEELPMTFKIVLEPHIEVNSKFNPKKASGMPLEPT 460

QY 380 NLSMILVAPDGLIVLHFKDIPNHDHFSVAIFTHKILPPDPSGVVCSVNVAGWVKPFNI 439

DB 361 SEEMTLVRPDCITVQPNDFNHDHFSVAIFTHKILPPDPSGVVCSVNVAGWVKPFNI 420

QY 440 SVKVLKPKLNAPNVDTGHNFAVINISSEPEFG-PKSCDKTHICP----- 484

DB 421 SVKVLKPKLNAPNVDTGHNFAVINISSEPEFG-PKSCDKTHICP----- 480

QY 485 -----PCPAPPELLGGP-----SVFLPPPKPKDTLMISRTIP-EV 516

DB 481 VTLNVLPTDYELGVQVLRPPGEGEHPVPRPFTTASIGLPPRGLSLIPKQSQTALNL 540

QY 517 TCVVVDVSHEDHPRKFNWYDGVFVHNAKIKPRFGQYNSIYRVSVLT-VLHGLWLNCKE 575

DB 541 TWQPIETSSD-----EFY---VEVERWQOTRSDQCN--IKVPGNLTSLVLLNLLFREQ 590

QY 576 YKCVSNKALPAPLETKYAKQO-PPREPQVYT-----LPPSREEM-TKNVSLTCLVK-- 627

DB 591 YSVRA-----PVNTKAGPFWSEFPAMVLSLOPPENIKIINIVTALVSWI 640

QY 628 ---GFYPSDIAVENESNCPENNY 648

DB 641 IVDGYSTSSIIIRYKVKQKNEQHH 664

## RESULT 9

US-08-484-438-10

: Sequence 10, Application US/08484438

: Patent No. 5811098

: Patent No. 5811098 5780031

: GENERAL INFORMATION:

: APPLICANT: Maisonnier, Gregory D.

: APPLICANT: Culouscou, Jean-Michel

: APPLICANT: Shoyab, Mohammed

: APPLICANT: Siedall, Clay B.

: APPLICANT: Hollistr m, Ingegerd

: APPLICANT: Hollistr m, Karl R.



;; TITLE OF INVENTION: HEP4 HUMAN RECEPTOR TYROSINE KINASE

;; NUMBER OF SEQUENCES: 42

;; CORRESPONDENCE ADDRESS:

;; ADDRESSEE: Pennie & Edmonds

;; STREET: 1155 Avenue of the Americas

;; CITY: New York

;; STATE: New York

;; COUNTRY: U.S.A.

;; ZIP: 10036-2711

;; COMPUTER READABLE FORM:

;; MEDIUM TYPE: Floppy disk

;; OPERATING SYSTEM: PC-DOS/MS-DOS

;; SOFTWARE: Patentin Release #1.0, Version #1.25

;; CURRENT APPLICATION DATA:

;; APPLICATION NUMBER: US/08/484,438

;; FILING DATE: 07-JUN-1995

;; CLASSIFICATION: 530

;; PRIOR APPLICATION DATA:

;; APPLICATION NUMBER: 08/323,442

;; FILING DATE: 14-OCT-1994

;; APPLICATION NUMBER: US 08/150,704

;; FILING DATE: 10-NOV-1993

;; CLASSIFICATION: 530

;; PRIOR APPLICATION DATA:

;; APPLICATION NUMBER: US 07/981,165

;; FILING DATE: 24-NOV-1992

;; CLASSIFICATION: 530

;; ATTORNEY/AGENT INFORMATION:

;; NAME: Misrock, S. Leslie

;; REGISTRATION NUMBER: 18,872

;; REFERENCE/DOCKET NUMBER: 5624-230

;; TELEPHONE: (212) 790 9090

;; TELEFAX: (212) 869-8864/9741

;; TELEX: 66141 PENNIE

;; INFORMATION FOR SEQ ID NO: 10:

;; SEQUENCE CHARACTERISTICS:

;; LENGTH: 911 amino acids

;; TYPE: amino acid

;; STRANDEDNESS: unknown

;; TOPOLOGY: unknown

;; MOLECULE TYPE: protein

;; US-08-484-438-10

Query Match 33.9%; Score 1311; DB 2; Length 911;

Best Local Similarity 51.2%; Pred No 16095;

Matches 288; Conservative 27; Mismatches 64; Indels 184; Gaps 21.

```

QY 198 LFTSAFTRLIVR-----RCEACKWGHCNHLDTA--CMNNG-----VCHP 235
DB 477 LFTSINGRIVRNKRAKNCNTAI--GMVCHLSSGCGGPGHDDQCLGKRRYSRGHCIC 534
QY 236 D TGE-----CI CPGFGMGRTCEK-----ACEHTFGRTIC 264
DB 535 SCNIYDGRFFRFGNCSGVCHGAD-----QCKKMFVGLIHCCHGPGNCKCKSHKICPN 589
QY 265 KERC SQGCKCKSVVGLHHPVGCSTATCKWGLQCNKACHPGHYGPKLPYS--CNNGF 321
DB 590 VEKGPQGLQANSEFTKYALP-----DPEHP-----CHENLTGPNMPT 629
QY 322 MDRFQFCGLCSGQWGLQCHRGIPRMTPKIVDLPDHLFVNSSCKFNPICKASCWPLPTNF 381
DB 630 SHD-----CIYYP-WTC-----HSTLPQ-----HPKAVKA-----LQFP 658
QY 382 EMTLVKPDGTVLHPKDFNHTDHFSAIFTIIRILPDDSGWVGSVNTVAGMVEKENISV 441
DB 659 --RLVGPD-----PFGCARPANTFLDPE----- 679
QY 442 KVLPKPLNAPNVIDIGHNFAVINISHPYFGPKSCDKTHTCPCHPAPLGLGPSVFLFP 501
DB 660 -----EPKSCDKTHTCPCHPAPLGLGPSVFLFP 708

```

```

QY 502 PKPKDILMISPTPEVTCVVVGVSHEDPEVKFNWYVDRVVEVHNAITKPPPEEYNSITPVVS 561
DB 709 PKPKDILMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVFVHNAITKPPPEEYNSITPVVS 768
QY 562 VLTVLHJLWLNKEYKCKVSNKALPAPIEKTSIKAPKAPPPQVYVTLFSPSEETKRVVS 621
DB 769 VLTVLHJLWLNKEYKCKVSNKALPAPIEKTSIKAKGQGFREFQVYVTLFSPSEETKRVVS 828
QY 622 LILVLKGFYPSDIAVWESNGQFNENNYKIIPVLLSDSGSFYLSKGLVDKSKWQGNVFS 681
DB 829 LILVLKGFYPSDIAVWESNGQFNENNYKIIPVLLSDSGSFYLSKGLVDKSKWQGNVFS 888
QY 682 GSVMEALHNHYTQKSLSLSPCK 704
DB 889 GSVMEALHNHYTQKSLSLSPCK 911

```

RESULT 10

US-08-385-229-4

;; Sequence 4, Application US/08385229

;; Patent No. 5605690

;; GENERAL INFORMATION:

;; APPLICANT: Jacobs, Cindy A.

;; APPLICANT: Smith, Craig A.

;; TITLE OF INVENTION: Method of Treating TNF-Dependent

;; TITLE OF INVENTION: Inflammation Using Tumor Necrosis Factor Antagonists

;; NUMBER OF SEQUENCES: 5

;; CORRESPONDENCE ADDRESS:

;; ADDRESSEE: Immunex Corporation

;; STREET: 51 Duversity Street

;; CITY: Seattle

;; STATE: Washington

;; COUNTRY: U.S.A.

;; ZIP: 98101

;; COMPUTER READABLE FORM:

;; MEDIUM TYPE: Floppy disk

;; OPERATING SYSTEM: PC-DOS/MS-DOS

;; SOFTWARE: Patentin Release #1.0, Version #1.25

;; CURRENT APPLICATION DATA:

;; APPLICATION NUMBER: US/08/385,229

;; FILING DATE:

;; CLASSIFICATION: 435

;; PRIOR APPLICATION DATA:

;; APPLICATION NUMBER: US/07/946,236

;; FILING DATE:

;; ATTORNEY/AGENT INFORMATION:

;; NAME: Wight, Christopher L.

;; REGISTRATION NUMBER: 31,690

;; REFERENCE/DOCKET NUMBER: 2503

;; TELECOMMUNICATION INFORMATION:

;; TELEPHONE: (206) 587-0430

;; TELEFAX: (206) 587-0606

;; INFORMATION FOR SEQ ID NO: 4:

;; SEQUENCE CHARACTERISTICS:

;; LENGTH: 518 amino acids

;; TYPE: amino acid

;; TOPOLOGY: linear

;; MOLECULE TYPE: protein

;; US-08-385-229-4

Query Match 33.6%; Score 1300; DB 1; Length 518;

Best Local Similarity 59.0%; Pred. No. 5,7695;

Matches 285; Conservative 29; Mismatches 105; Indels 72; Gaps 17;

```

QY 248 GPTTEKAKELHTFGRTTKEFGSGEGSKSYVGLPLPVPVLSAIGWGLQVNEACHPGFY 307
DB 66 GSTCHLRYHYDGTAGCCCKSCSPQGHAK--VPC-----TKISDVC-DSECSIY 112
QY 308 GP-----DKLPGSTNNFE--MTTFEG--TLPKRWGLQVNECHPPTPTPKIVD 354
US-08-385-229-4

```

DB 113 TOLWNVPECLSGKSSQVETQATREQNRIITCFPFWYALSKQEGRLCAP---- 168  
QY 355 LPDHLFVNSCKFNPKCKASQWPLNEDMTLV-KDGLIVLHPKDFNHTDHSVAITFHR 413  
DB 169 -----LRKCP-----GFGVAPRGTFTSIVVCKP-----CAPGTFSNTS-STDICRPHO 212  
QY 414 1-----LPPDSGV-WVCSVNT-----VAGMVEKPFNISVK-----VLPKPAPNPNVIDTG 457  
DB 213 1CNVVAIPGNSMAVATSTSPFSMAZAVHLPQVPSIRSHQIQTPTPESTAPST---- 268  
QY 458 HNFVAVINISSHPYFC--EPKSCDKTHICPPCPAPFELAGGSVFELPPKPKDILMISRT 513  
DB 269 SFLLPMKPSPAEASTGSEKSTKTHICPPCPAPFELAGGSVFELPPKPKDILMISRT 327  
QY 514 PRVTCVVVDSHEDPEVFNWYVDCVGVHNAKTKPREQYNSYRVVSVTLVLHQWMLNG 573  
DB 328 PRVTCVVVDSHEDPEVFNWYVDCVGVHNAKTKPREQYNSYRVVSVTLVLHQWMLNG 387  
QY 574 KEYCKVSNKALPAPTEKTSKAKGQPREPOVYTLPPSPPEMTKNQVSLTCLVKGEYPSD 633  
DB 388 KDKCKVSNKALPAPMOKTISKAKGQPREPOVYTLPPSRDELTKNOVSLTCLVKGEYPRH 447  
QY 634 IAVVESNGQPNNTKLTTPVLDSDGSFFLYSKLTVDKSRWQGNVFCSCVMHEALHNY 693  
DB 448 IAVVESNGQPNNTKLTTPVLDSDGSFFLYSKLTVDKSRWQGNVFCSCVMHEALHNY 507  
QY 694 1OKSLSLSPGK 704  
DB 508 1OKSLSLSPGK 518

## RESULT 11

US-08-784-512-1

Sequence 1, Application US/08/243013

Patent No. 5639597

## GENERAL INFORMATION:

APPLICANT: Laulier, Leander

APPLICANT: Zellmeissel, Gerd

APPLICANT: Quendo, Patricia

TITLE OF INVENTION: Cell free Receptor Binding Assays, the

TITLE OF INVENTION: Production and Use Thereof

NUMBER OF SEQUENCES: 6

CORRESPONDENCE ADDRESS:

ADDRESSEE: Elmouan, Henderson, Farabow, Garrett &amp;

ADDRESSEE: Dumer

STREET: 1460 I Street, N.W.

CITY: Washington

STATE: D.C.

COUNTRY: USA

ZIP: 20005-3315

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/243,010

FILING DATE: 13-MAY-1994

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 97/798,564

FILING DATE: 26-NOV-1991

APPLICATION NUMBER: DE P 40 37 847.3

FILING DATE: 28-NOV-1990

ATTORNEY/AGENT INFORMATION:

NAME: Einadly, Carol P.

REGISTRATION NUMBER: 32,220

REFERENCE/INFORMATION: 92481-1132-00000

TELECOMMUNICATION INFORMATION:

TELEPHONE: 202-408-4000

TELEFAX: 202-408-4400

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 486 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-243-010-1

Query Match 33.68; Score 1298.5; DB 1; Length 486;  
Residual Similarity 58.98; Pred. No. 6,90-95;  
Matches 288; Conservative 26; Mismatches 104; Indels 71; Gaps 18;

QY 248 GRVTEKACELHIFGRICKERGSQVGVKLPDPYGGSCATGKWKGLQFNEACHRGY 407  
DB 37 GSICELREYVYQTAUMCGSKSPQHIAK--VFG-----IKSDIWC-DSCELSY 84  
QY 308 GP-----DCKLRSCNNCP--MCDPRG--CLSPGCGGLQFEPGIPRMTKIVD 454  
DB 84 TOLWNVPECLSGKSSQVETQATREQNRIITCFPFWYALSKQEGRLCAP---- 169  
QY 355 LPDHLFVNSCKFNPKCKASQWPLNEDMTLV-KDGLIVLHPKDFNHTDHSVAITFHR 413  
DB 169 -----LRKCP-----GFGVAPRGTFTSIVVCKP-----CAPGTFSNTS-STDICRPHO 212  
QY 414 1-----LPPDSGV-WVCSVNT-----VAGMVEKPFNISVK-----VLPKPAPNPNVIDTG 457  
DB 213 1CNVVAIPGNSMAVATSTSPFSMAZAVHLPQVPSIRSHQIQTPTPESTAPST---- 268  
QY 458 HNFVAVINISSHPYFC--EPKSCDKTHICPPCPAPFELAGGSVFELPPKPKDILMISRT 513  
DB 269 SFLLPMKPSPAEASTGSEKSTKTHICPPCPAPFELAGGSVFELPPKPKDILMISRT 327  
QY 514 PRVTCVVVDSHEDPEVFNWYVDCVGVHNAKTKPREQYNSYRVVSVTLVLHQWMLNG 573  
DB 328 PRVTCVVVDSHEDPEVFNWYVDCVGVHNAKTKPREQYNSYRVVSVTLVLHQWMLNG 387  
QY 574 KEYCKVSNKALPAPTEKTSKAKGQPREPOVYTLPPSPPEMTKNQVSLTCLVKGEYPSD 633  
DB 388 KDKCKVSNKALPAPMOKTISKAKGQPREPOVYTLPPSRDELTKNOVSLTCLVKGEYPRH 447  
QY 634 IAVVESNGQPNNTKLTTPVLDSDGSFFLYSKLTVDKSRWQGNVFCSCVMHEALHNY 693  
DB 448 IAVVESNGQPNNTKLTTPVLDSDGSFFLYSKLTVDKSRWQGNVFCSCVMHEALHNY 507  
QY 694 1OKSLSLSPGK 704  
DB 508 1OKSLSLSPGK 518

## RESULT 12

US-08-784-512-3

Sequence 3, Application US/08784512

Patent No. 5872209

## GENERAL INFORMATION:

APPLICANT: BARTNIK, Eckart

APPLICANT: EIDENMUELLER, Bernd

APPLICANT: BUETTNER, Frank

APPLICANT: CATHERSON, Bruce

APPLICANT: HUGHES, Claire

TITLE OF INVENTION: An artificial recombinant substrate (RAGG 1)

TITLE OF INVENTION: and native aggrecan to study the proteolytic activity of

TITLE OF INVENTION: "Aggrecanase" in cell culture systems

NUMBER OF SEQUENCES: 4

CORRESPONDENCE ADDRESS:

ADDRESSEE: Foley &amp; Jardner

STREET: Suite 500, 3000 K Street, N.W.

CITY: Washington, D.C.

COUNTRY: USA

ZIP: 20007-5109

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.25

: CURRENT APPLICATION DATA:  
 : APPLICATION NUMBER: US/08/784,512  
 : FILING DATE: 17-JAN-1997  
 : PRIOR APPLICATION DATA:  
 : APPLICATION NUMBER: EP 96100682.2  
 : FILING DATE: 18-JAN-1996  
 : ATTORNEY/AGENT INFORMATION:  
 : NAME: GRANADOS, Patricia D.  
 : REGISTRATION NUMBER: 33,683  
 : REFERENCE/DOCKET NUMBER: 18748/311  
 : TELECOMMUNICATION INFORMATION:  
 : TELEPHONE: (202)572-5300  
 : TELEFAX: (202)672-5399  
 : TELEX: 904136  
 : INFORMATION FOR SEQ ID NO: 3:  
 : SEQUENCE CHARACTERISTICS:  
 : LENGTH: 396 amino acids  
 : TYPE: amino acid  
 : STRANDEDNESS: single  
 : TOPOLOGY: linear  
 : MOLECULE TYPE: protein  
 : FEATURE:  
 : NAME/KEY: Protein  
 : LOCATION: 1..396  
 : US-08-784-512-3

Query Match 33.1%; Score 1279.5; DB 2; Length 396;  
 Best Local Similarity 68.4%; Pred. No. 1.7e-93;  
 Matches 255; Conservative 21; Mismatches 62; Indels 35; Gaps 7;  
 QY 353 VDLPHIEVNSKFNPKICKASGP---LPTNEEMTLVKGDTGL-----HPKDFNH 400  
 Db 38 VDIPEFEGVGGEEDITVQVTWPMELPLRNITEGEARGSVILTVKPIFVSPSLEP 97  
 QY 401 IDHFSVAIFTIHRILPPDSGVWV-CSVNTVAGMVKPKFNISVKVLKPKLNAPNVIDTGHN 459  
 Db 98 EEPFTFA-----PEIGATAFAVENETGEATRFWG-----FPIPGIGPATATFSED 143  
 QY 460 FA--VINISSEPF-----GPKSKDTHICPPCPAPHELGGPSVFLPPKPKDTLMS 511  
 Db 144 LVQVTVAPGQPHLPGGDPEPKSCUKTHCPHCPAPHELGGPSVFLPPKPKDTLMS 203  
 QY 512 RPEVTCVVDVSHEDPEVKFNWYDGVGVHNAKTPREEQYNSTYRVSVLTLHQDWL 571  
 Db 204 RTPEVTCVVDVSHEDPEVKFNWYDGVGVHNAKTPREEQYNSTYRVSVLTLHQDWL 263  
 QY 572 NGKEYCKVSNKALPAPIEKTISKAKQPREPOVYTLPPSRPEMTKNQVSLTCLVKGFP 631  
 Db 264 NGKEYCKVSNKALPAPIEKTISKAKQPREPOVYTLPPSRDELTKNQVSLTCLVKGFP 323  
 QY 632 SDIAVWESNGOPNNYKTPPVLDSGSPFLYSLKLIIVKSKWQOCNVFSCVMHEALHN 691  
 Db 384 HYTKSLSLSPGK 396

RESULT 13  
 US-09-176-228-3  
 : Sequence 3, Application US/09176228  
 : Patent No. 6180334  
 : GENERAL INFORMATION:  
 : APPLICANT: RABINIK, Eckart  
 : APPLICANT: EIDENMUELLER, Bernd  
 : APPLICANT: BUETINER, Frank  
 : APPLICANT: CATERSON, Bruce  
 : APPLICANT: HUGHES, Clare  
 : TITLE OF INVENTION: An artificial recombinant substrate (rAG 1)  
 : TITLE OF INVENTION: And native aggracan to study the proteolytic activity of  
 : TITLE OF INVENTION: "Aggracanase" in cell culture systems

: NUMBER OF SEQUENCES: 4  
 : CORRESPONDENCE ADDRESS:  
 : ADDRESSEE: Foley & Lardner  
 : STREET: Suite 500, 3000 K Street, N.W.  
 : CITY: Washington, D.C.  
 : COUNTRY: USA  
 : ZIP: 20007-5109  
 : COMPUTER READABLE FORM:  
 : MEDIUM TYPE: Floppy disk  
 : COMPUTER: IBM PC compatible  
 : OPERATING SYSTEM: PC-DOS/MS-DOS  
 : SOFTWARE: PatentIn Release #1.0, Version #1.25  
 : CURRENT APPLICATION DATA:  
 : APPLICATION NUMBER: US/09/176,228  
 : FILING DATE:  
 : PRIOR APPLICATION DATA:  
 : APPLICATION NUMBER: US/08/784,512  
 : FILING DATE: 17-JAN-1997  
 : APPLICATION NUMBER: EP 96100682.2  
 : FILING DATE: 18-JAN-1996  
 : ATTORNEY/AGENT INFORMATION:  
 : NAME: GRANADOS, Patricia D.  
 : REGISTRATION NUMBER: 33,683  
 : REFERENCE/DOCKET NUMBER: 18748/311  
 : TELECOMMUNICATION INFORMATION:  
 : TELEPHONE: (202)672-5300  
 : TELEFAX: (202)672-5399  
 : TELEX: 904136  
 : INFORMATION FOR SEQ ID NO: 3:  
 : SEQUENCE CHARACTERISTICS:  
 : LENGTH: 396 amino acids  
 : TYPE: amino acid  
 : STRANDEDNESS: single  
 : TOPOLOGY: linear  
 : MOLECULE TYPE: protein  
 : FEATURE:  
 : NAME/KEY: Protein  
 : LOCATION: 1..396  
 : US-09-176-228-3

Query Match 33.1%; Score 1279.5; DB 4; Length 396;  
 Best Local Similarity 68.4%; Pred. No. 1.7e-93;  
 Matches 255; Conservative 21; Mismatches 62; Indels 35; Gaps 7;  
 QY 353 VDLPHIEVNSKFNPKICKASGP---LPTNEEMTLVKGDTGL-----HPKDFNH 400  
 Db 38 VDIPEFEGVGGEEDITVQVTWPMELPLRNITEGEARGSVILTVKPIFVSPSLEP 97  
 QY 401 IDHFSVAIFTIHRILPPDSGVWV-CSVNTVAGMVKPKFNISVKVLKPKLNAPNVIDTGHN 459  
 Db 98 EEPFTFA-----PEIGATAFAVENETGEATRFWG-----FPIPGIGPATATFSED 143  
 QY 460 FA--VINISSEPF-----GPKSKDTHICPPCPAPHELGGPSVFLPPKPKDTLMS 511  
 Db 144 LVQVTVAPGQPHLPGGDPEPKSCUKTHCPHCPAPHELGGPSVFLPPKPKDTLMS 203  
 QY 512 RPEVTCVVDVSHEDPEVKFNWYDGVGVHNAKTPREEQYNSTYRVSVLTLHQDWL 571  
 Db 204 RTPEVTCVVDVSHEDPEVKFNWYDGVGVHNAKTPREEQYNSTYRVSVLTLHQDWL 263  
 QY 572 NGKEYCKVSNKALPAPIEKTISKAKQPREPOVYTLPPSRPEMTKNQVSLTCLVKGFP 631  
 Db 264 NGKEYCKVSNKALPAPIEKTISKAKQPREPOVYTLPPSRDELTKNQVSLTCLVKGFP 323  
 QY 632 SDIAVWESNGOPNNYKTPPVLDSGSPFLYSLKLIIVKSKWQOCNVFSCVMHEALHN 691  
 Db 384 HYTKSLSLSPGK 396

```

RESULT 14
US-08-097-827-11
: Sequence 11, Application US/08/97827
: GENERAL INFORMATION:
: APPLICANT: Baum, Peter
:               Goodwin, Ray
:               Fanslow, William
:               Gayle, Richard
: TITLE OF INVENTION: Novel Cytokine which is a ligand for
:               OX40
: NUMBER OF SEQUENCES: 13
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Immunex Corporation
: STREET: 51 University Street
: CITY: Seattle
: STATE: WA
: COUNTRY: USA
: ZIP: 98101
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent In Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/097.827
: FILING DATE: 23-Jul-1993
: CLASSIFICATION: <Unknown>
: ATTORNEY/AGENT INFORMATION:
: NAME: Perkins, Patricia A.
: REGISTRATION NUMBER: 34,693
: REFERENCE/DOCKET NUMBER: 2806
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 206-587-0730
: INFORMATION FOR SEQ ID NO: 11:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 438 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: SEQUENCE DESCRIPTION: SEQ ID NO: 11:
US 08 097-827 11

```

```

Query Match      33.0%; Score 1276.5; DB 1; Length 438;
Best Local Similarity 55.5%; Pred. No. 3,3e-93;
Matches 274; Conservative 23; Mismatches 88; Indels 109; Gaps 14;

QY 247 MGRTEKA-CELITF--GRTCKEKS--SGEGKSYVF--IPFYAS--AT--SNKGLA--NEA-- 301
DB 18 LGVTARRLNCVKRHVHPSCHKCRCKCPCHGMVN-----KCDHTRD 57

QY 402 --CHP--GPGV-----DCKLRCSCNNGEMCDREQGL-----CSPGWGLQERE 343
DB 58 TLCHVCEGFTFNAVNYDICKAGLCQCNHRSGLKQK--IPTQDIVCKRGTQ----- 110

QY 344 GIPRMPKIVLVDLPHLEVNSCKFNI-----ICKASQWPL--PINEMLIVKPGT 391
DB 111 --PRQUSYKLVLCV--CP--HFSFGNNQA--KPWINCI--LSKQTRHFASSDGLIAV--EDRS 168

QY 492 VLPKPNHTDHFSAVETIHTIPPLPSGWWG--SVNTVA--MVEKPEPI--SVFVLPKPLNA 450
DB 169 LL-----ATLWETQRTF-----HPTVQSTVWVPRTSPT 199

QY 451 PNVIDTGNFAVINISSEFYCEPKSKDTHTTQPCAPPELLIGPSVFLPPPKDPLMI 510
DB 200 P-----STETLEVPSTKTHFTPTCTAPEAESAPSELPPEPKPTLMI 244

QY 511 SKRPVTCVVIVDSHEDDEVFNNVYDCVGVHNAKTRPREYANSTYKVSVLTVLHQW 570
DB 245 SKRPVTCVVIVDSHEDDEVFNNVYDCVGVHNAKTRPREYANSTYKVSVLTVLHQW 564

QY 571 LNCKEYKCVSNKALPAPIETKTISKAKGQPPPEPVYTLPPSPDEMTKNVSLTCLVKGFY 630

```

```

DB 305 LNCKEYKCVSNKALPAPIETKTISKAKGQPPPEPVYTLPPSPDEMTKNVSLTCLVKGFY 664
QY 631 PSDIAVENHSCNCPENNYKTTTPPVLDSDGSPFLYSKLTVDKSHRQCCNVFSCSYMHIALH 690
DB 365 PSDIAVENHSCNCPENNYKTTTPPVLDSDGSPFLYSKLTVDKSHRQCCNVFSCSYMHIALH 424
QY 691 NHYTOKSLSLSPGK 704
DB 425 NHYTOKSLSLSPGK 438

RESULT 15
US-08-494-574-11
: Sequence 11, Application US/08/494574
: Patent No. 5783665
: GENERAL INFORMATION:
: APPLICANT: Baum, Peter
: APPLICANT: Goodwin, Ray
: APPLICANT: Fanslow, William
: APPLICANT: Gayle, Richard
: TITLE OF INVENTION: No. 5783665cl Cytokine which is a ligand for
: TITLE OF INVENTION: OX40
: NUMBER OF SEQUENCES: 13
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Immunex Corporation
: STREET: 51 University Street
: CITY: Seattle
: STATE: WA
: COUNTRY: USA
: ZIP: 98101
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent In Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/494,574
: FILING DATE: 22-JUN-1995
: CLASSIFICATION: 530
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US/08/097,827
: FILING DATE: 23-JUL-1993
: ATTORNEY/AGENT INFORMATION:
: NAME: Perkins, Patricia A.
: REGISTRATION NUMBER: 34,693
: REFERENCE/DOCKET NUMBER: 2806
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 206-587-0730
: INFORMATION FOR SEQ ID NO: 11:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 438 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
US-08-494-574-11

```

```

Query Match      33.0%; Score 1276.5; DB 1; Length 438;
Best Local Similarity 55.5%; Pred. No. 3,3e-93;
Matches 274; Conservative 23; Mismatches 88; Indels 109; Gaps 14;

QY 247 MGRTEKA-CELITF--GRTCKEKS--SGEGKSYVF--IPFYAS--AT--SNKGLA--NEA-- 401
DB 18 LGVTARRLNCVKRHVHPSCHKCRCKCPCHGMVN-----KCDHTRD 57

QY 302 --CHP--GPGV-----DCKLRCSCNNGEMCDREQGL-----CSPGWGLQERE 443
DB 58 TLCHVCEGFTFNAVNYDICKAGLCQCNHRSGLKQK--IPTQDIVCKRGTQ----- 110

QY 344 GIPRMPKIVLVDLPHLEVNSCKFNI-----ICKASQWPL--PINEMLIVKPGT 491
DB 111 --PRQUSYKLVLCV--CP--HFSFGNNQA--KPWINCI--LSKQTRHFASSDGLIAV--EDRS 168

```

```

QY 392 VLHFKDFMHTHFSVALETHIRILPDSGVWVGSVNTVAGMVEKPENI-SVKVLKPLNA 450
Db 169 LL-----ATLLWETQRIPTF-----RPTTVOSTTIVWPTSMI 199
QY 451 PNVIDTGHNEAVINISSEPYFGEKSCDKTHTCPGAPAPPELLGGPSVFLFPPKPKDTLMI 510
Db 200 P-----STPTLVEFRSCDKTHTCPGAPAEAGAPSVFLFPPKPKDTLMI 244
QY 511 SRTPETVCVVVDVSHEDPEVKFNWYVDGVEVHNAKIKPREEOYNSTYRVVSVLTVLHQDW 570
Db 245 SRTPETVCVVVDVSHEDPEVKFNWYVDGVEVHNAKIKPREEOYNSTYRVVSVLTVLHQDW 304
QY 571 LNKKEYCKVSNKALPAPIEKTISKAKGQPREPKQVYTLPPSRREEMTKNAVSLTCLVKGFY 630
Db 305 INCKEYCKVSNKALPAPIEKTISKAKGQPREPKQVYTLPPSRREEMTKNAVSLTCLVKGFY 364
QY 631 PSDIAVEWESNGGPPENNYKTTTPVLDSDGSFFLYSKLTVDKSRWQQGNVFSQSVMEALH 690
Db 365 PSDIAVEWESNGGPPENNYKTTTPVLDSDGSFFLYSKLTVDKSRWQQGNVFSQSVMEALH 424
QY 691 NHYTKQSLSLSPGK 704
Db 425 NHYTKQSLSLSPGK 438

```

Search completed: September 23, 2002, 22:10:22  
Job time: 423 sec



GenCore version 4.5  
Copyright (c) 1994 - 2000 Compugen Ltd

OM protein - protein search, using SW model

Run on: September 23, 2002, 22:08:44, Search time: 45.63 seconds  
(without alignments)  
1485.765 Million cell updates/sec

Title: US-09-733-764-2  
Perfect score: 3868  
Sequence: 1 MDSIASIVGVSVSLISGVTEGAMDLITINSIPVSVAPTSITCIASGWRPHEPTIGRD 60  
Scoring table: BLOSOM62  
Gapop 10.0, Gapext 0.5

Searched: 283136 seqs, 36089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match: 94

Maximum Match: 100%

Listing first 45 summaries

Database:

1: p1r1: \*  
2: p1r2: \*  
3: p1r3: \*  
4: p1r4: \*

Prod No is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Posut	No.	Score	Query	Match	Length	DE	ID	Description
1	2623	67.8	1124	1	158388			protein-tyrosine k
2	2458	63.5	1125	1	57842			protein-tyrosine k
3	2432.5	62.9	1123	2	154237			protein-tyrosine k
4	2363	61.1	1125	1	106771			protein-tyrosine k
5	1271	32.9	974	2	874139			protein-tyrosine k
6	1263	32.6	930	1	3100			Ig gamma-1 chain C
7	1243	32.3	955	4	833866			Ig gamma-1 chain C
8	1204	31.1	234	2	175207			Ig gamma-1 chain C
9	1183	30.4	377	2	A21511			Ig gamma-1 chain C
10	1181	30.5	377	2	A60764			Ig gamma-3 chain C
11	1158	29.9	289	1	338701			Ig gamma-3 heavy c
12	1154	29.8	326	1	3200			Ig gamma-3 heavy c
13	1144	29.6	327	1	64100			Ig gamma-4 chain C
14	932	24.1	323	1	64100			Ig gamma-4 chain C
15	915	23.7	328	2	147159			Ig gamma chain C r
16	910.5	23.5	328	2	147160			Ig gamma 2a chain
17	906.5	23.4	277	2	147162			Ig gamma 2b chain
18	893.5	23.1	329	1	G26F			Ig gamma 4 chain C
19	892	23.1	928	2	147159			Ig gamma 2 chain C
20	883	22.8	928	2	147161			Ig gamma 3 chain C
21	882.5	22.8	1136	1	S57845			protein-tyrosine k
22	874.5	22.6	1138	1	S21066			protein-tyrosine k
23	866	22.4	470	2	S22080			Ig heavy chain pre
24	857.5	22.2	325	2	3100			Ig gamma-3 chain C
25	851.5	22.0	333	2	PS0318			Ig gamma-2b chain
26	849.5	22.0	308	2	338754			Ig heavy chain C r
27	849.5	22.0	472	2	S31459			Ig gamma-1 chain r
28	849.5	22.0	472	2	S31459			monoclonal antibody
29	848	21.9	444	2	104436			

## ALIGNMENTS

RESULT 1

158388

Protein tyrosine kinase (EC 2.7.1.112), receptor type tek precursor - human  
C:Species: Homo sapiens (man)  
C:Date: 02-Jul-1992 #sequence\_revision: 02 Jul 1992 #text\_change: 16-Jul-1999  
C:Accession: 158388  
R:Ziegler, S.F.; Bird, T.A.; Schneider, J.A.; Schooley, K.A.; Baum, P.R.  
Oncogene 8, 663-670, 1993  
A:Title: Molecular cloning and characterization of a novel receptor protein tyrosine  
A:Reference number: 158388; MUID:93173509  
A:Accession: 158388  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-1124 <RES>  
A:Cross references: GR:106133; NID:9292823, PID:9292823, PID:9292824  
C:Genetics:  
A:Gene: GDB:TKK  
A:Cross references: GDB:344185; OMIM:600221  
A:Map position: 9p21-9p21  
C:Function:

C:Description: Catalyzes the phosphorylation of a peptidyl tyrosine residue by ATP  
C:Superfamily: Protein tyrosine kinase, receptor type tek, EGF homology, fibronectin  
C:Keywords: ATP, autophosphorylation, duplication, glycoprotein, phosphoprotein, phos  
E:222/Domain: Signal sequence status predicted <SIS>  
E:23-124/Fibronectin protein tyrosine kinase, type III repeat homology <FN3>  
E:37-104/Domain: immunoglobulin homology <IM1>  
E:135-137/Region: Cell attachment (R-G-D) motif  
E:211-251/Domain: EGF homology <E21>  
E:255-298/Domain: EGF homology <E22>  
E:302-340/Domain: EGF homology <E23>  
E:364-426/Domain: immunoglobulin homology <IM2>  
E:447-527/Domain: fibronectin type III repeat homology <FN1A>  
E:542-625/Domain: fibronectin type III repeat homology <FN1B>  
E:638-720/Domain: fibronectin type III repeat homology <FN1C>  
E:722-772/Domain: transmembrane status predicted <TM>  
E:822-1097/Domain: protein kinase homology <KIN>  
E:830-838/Region: protein kinase ATP-binding motif  
E:140-158/Region: 438,454,560,566,645,651/Binding sites: carbohydrate (Asn); (covalent) #st  
E:855,872,964/Active site: Lys, Glu, Asp #status predicted

Query Match 67.8% Score 2523 26 27 Length 1144

Best Local Similarity 77.1% Pred. No. 245-139;

Matches 501, Conservative 23, Mismatches 44, Indels 82, Gaps 7;

QY 1 MDSIASIVGVSVSLISGVTEGAMDLITINSIPVSVAPTSITCIASGWRPHEPTIGRD 60

|||||  
1 MDSIASIVGVSVSLISGVTEGAMDLITINSIPVSVAPTSITCIASGWRPHEPTIGRD 60

QY 61 FEALMRHLEFLEVLTVFTEWAFVWVWPKERASEFNAYFESGVFSEAEIEEMKEL 120

|||||  
61 FEALMRHLEFLEVLTVFTEWAFVWVWPKERASEFNAYFESGVFSEAEIEEMKEL 120

QY 61 FEALMRHLEFLEVLTVFTEWAFVWVWPKERASEFNAYFESGVFSEAEIEEMKEL 120

|||||  
61 FEALMRHLEFLEVLTVFTEWAFVWVWPKERASEFNAYFESGVFSEAEIEEMKEL 120

```

121 QASFLPATITMVDKGNVNSKAVIETKEEAVIYKNGSETHSVPRHEVDILEVHLPH 180
|||||
121 QASFLPATITMVDKGNVNSKAVIETKEEAVIYKNGSETHSVPRHEVDILEVHLPH 180
|||||
181 AOHQAGVYSARYIGNNIETSAFTPIIVPRCEAKWPEKPEKTHETACMNNVWHEGDTGC 240
|||||
181 AOHQAGVYSARYIGNNIETSAFTPIIVPRCEAKWPEKPEKTHETACMNNVWHEGDTGC 240
|||||
241 ICPGPGMGTGCEKAVELHITGEGTCKEKEKSGAFKSKSYVFLPEGPGYSCATGKGLQNE 300
|||||
241 ICPGPGMGTGCEKAVELHITGEGTCKEKEKSGAFKSKSYVFLPEGPGYSCATGKGLQNE 300
|||||
301 ACHPGYCHDCKLRSCNNGEMCHREFORGLASQWGGIQQEREGIPRMTKIVLPLDHEIE 360
|||||
301 ACHPGYCHDCKLRSCNNGEMCHREFORGLASQWGGIQQEREGIPRMTKIVLPLDHEIE 360
|||||
361 VNSGKFNPICKASGWHPTNEMTAVKQIGVILHFKDNRHTDHSVAFTIHRILPDDSG 420
|||||
361 VNSGKFNPICKASGWHPTNEMTAVKQIGVILHFKDNRHTDHSVAFTIHRILPDDSG 420
|||||
421 VVWCSVNTVAGWVEKPEKNSVAVLPEKFNAPRVDTGHNFAVINISSEPEGEKSCOKT 480
|||||
421 VVWCSVNTVAGWVEKPEKNSVAVLPEKFNAPRVDTGHNFAVINISSEPEGEKSCOKT 480
|||||
481 HTCPGPAPELIGAGSVEFLPEPEKDTLMSIKTEPEVTCVVVDVSHEDPEVKNNYV 536
|||||
481 HTCPGPAPELIGAGSVEFLPEPEKDTLMSIKTEPEVTCVVVDVSHEDPEVKNNYV 536
|||||
536 KAKGQRPPEPVVTLDESEKEMTKNVSITGLVKRGVPS---DIAVEWESNGCPENN 647
|||||
536 KAKGQRPPEPVVTLDESEKEMTKNVSITGLVKRGVPS---DIAVEWESNGCPENN 647
|||||
539 ASIG-----LPPKCLSLIPKSOITLNTWQPIFPSSDDFYVEVEKRSVQMS 587
|||||
539 ASIG-----LPPKCLSLIPKSOITLNTWQPIFPSSDDFYVEVEKRSVQMS 587
|||||
RESULT 2
S57846
protein-tyrosine kinase (EC 2.7.1.112), receptor type tek precursor bovine
N:Alternate names: Has primigenius tarsus (cat10)
C:Species: Mus sp. (mouse)
C:Superfamily: ATP autophosphorylation duplication, glycoprotein phosphatase
C:Keywords: ATP autophosphorylation duplication, glycoprotein phosphatase
E:1-22/Domain: signal sequence #status predicted <SIG>
F:1125/Product: protein tyrosine kinase, receptor type tek #status predicted <MAT>
F:137-104/Domain: immunoglobulin homology <IM1>
F:135-137/Region: cell attachment (R-G-D) motif
F:211-251/Domain: EGF homology <EG1>
F:255-298/Domain: EGF homology <EG2>
F:302-340/Domain: EGF homology <EG3>
F:364-426/Domain: immunoglobulin homology <IM2>
F:447-527/Domain: fibronectin type III repeat homology <FN3>

```

```

F:542-626/Domain: fibronectin type III repeat homology <FN3B>
F:639-743/Domain: fibronectin type III repeat homology <FN3C>
F:753-773/Domain: transmembrane #status predicted <TM>
F:823-1130/Domain: protein kinase homology <PKIN>
F:831-839/Region: protein kinase ATP-binding motif
F:140,159,199,439,444,450,492,492/Binding site: catalytic site (Asn) (catalytic) #st
F:856,873,965/Active site: Lys, Glu, Asp #status predicted

Query Match 63.5%; Score 2458; DB 1; Length 1125;
Best Local Similarity 71.6%; Pred. No. 3,6e-140;
Matches 471; Conservative 35; Mismatches 76; Indels 82; Gaps 7;

QY 1 MSLASLVAGVSLSSGLIVGAMILLINSLPVSVAEISLGLASGWRHETITGRD 60
|||||
DB 1 MSLASLVAGVSLSSGLIVGAMILLINSLPVSVAEISLGLASGWRHETITGRD 60
|||||
QY 61 PPA: MNGCHDPIEVGIVWIPWAKKVVWVPEKASKINZAYEDSGVPEEAIRPIIMKMPQ 120
|||||
DB 61 PPA: MNGCHDPIEVGIVWIPWAKKVVWVPEKASKINZAYEDSGVPEEAIRPIIMKMPQ 120
|||||
QY 121 QASFLPATITMVDKGNVNSKAVIETKEEAVIYKNGSETHSVPRHEVDILEVHLPH 180
|||||
DB 121 QASFLPATITMVDKGNVNSKAVIETKEEAVIYKNGSETHSVPRHEVDILEVHLPH 180
|||||
QY 181 AOHQAGVYSARYIGNNIETSAFTPIIVPRCEAKWPEKPEKTHETACMNNVWHEGDTGC 240
|||||
DB 181 AOHQAGVYSARYIGNNIETSAFTPIIVPRCEAKWPEKPEKTHETACMNNVWHEGDTGC 240
|||||
QY 241 ICPGPGMGTGCEKAVELHITGEGTCKEKEKSGAFKSKSYVFLPEGPGYSCATGKGLQNE 300
|||||
DB 241 ICPGPGMGTGCEKAVELHITGEGTCKEKEKSGAFKSKSYVFLPEGPGYSCATGKGLQNE 300
|||||
QY 301 ACHPGYCHDCKLRSCNNGEMCHREFORGLASQWGGIQQEREGIPRMTKIVLPLDHEIE 360
|||||
DB 301 ACHPGYCHDCKLRSCNNGEMCHREFORGLASQWGGIQQEREGIPRMTKIVLPLDHEIE 360
|||||
QY 361 VNSGKFNPICKASGWHPTNEMTAVKQIGVILHFKDNRHTDHSVAFTIHRILPDDSG 420
|||||
DB 361 VNSGKFNPICKASGWHPTNEMTAVKQIGVILHFKDNRHTDHSVAFTIHRILPDDSG 420
|||||
QY 421 VVWCSVNTVAGWVEKPEKNSVAVLPEKFNAPRVDTGHNFAVINISSEPEGEKSCOKT 480
|||||
DB 421 VVWCSVNTVAGWVEKPEKNSVAVLPEKFNAPRVDTGHNFAVINISSEPEGEKSCOKT 480
|||||
QY 481 HTCPGPAPELIGAGSVEFLPEPEKDTLMSIKTEPEVTCVVVDVSHEDPEVKNNYV 536
|||||
DB 481 HTCPGPAPELIGAGSVEFLPEPEKDTLMSIKTEPEVTCVVVDVSHEDPEVKNNYV 536
|||||
QY 536 KAKGQRPPEPVVTLDESEKEMTKNVSITGLVKRGVPS---DIAVEWESNGCPENN 647
|||||
DB 536 KAKGQRPPEPVVTLDESEKEMTKNVSITGLVKRGVPS---DIAVEWESNGCPENN 647
|||||
QY 539 ASIG-----LPPKCLSLIPKSOITLNTWQPIFPSSDDFYVEVEKRSVQMS 587
|||||
DB 539 ASIG-----LPPKCLSLIPKSOITLNTWQPIFPSSDDFYVEVEKRSVQMS 587
|||||

RESULT 3
154237
protein tyrosine kinase (EC 2.7.1.112), receptor type tek precursor mouse
N:Alternate names: protein-tyrosine kinase receptor type tie2
C:Species: Mus sp. (mouse)
C:Date: 02-Aug-1996 #sequence revision 02-Aug-1996 #text change 21-Jan-2000
C:Accession: 154237
R:Kuntz, A.S., Stacker, S.A., Wilks, A.F.
Growth factors 9, 99-105, 1993
A:Title: tie2, a putative protein tyrosine kinase from a new class of cell surface re
A:Reference number: 154237; MUID:9403116
A:Accession: 154237
A:Status: preliminary, translated from cDNA, EMBL, DDBJ
A:Molecule type: mRNA
A:Residues: 1-1125 <RES>

```



A:Cross-references: QJ567951; NID:452873; EMD:AAH28663; EMB:4452974

C:Genetics:

A:Gene: tie2

C:Superfamily: protein-tyrosine kinase, receptor type III, EGF homology, fibronectin type III; autophosphorylation, duplication, glycoprotein, phosphoprotein, phospho

F:37-104/Domain: immunoglobulin homology <IMM>

F:211-251/Domain: EGF homology <EGF>

F:255-298/Domain: EGF homology <EGF>

F:302-340/Domain: fibronectin type III repeat homology <FN3>

F:540-624/Domain: fibronectin type III repeat homology <FN3>

F:820-1097/Domain: protein kinase homology <PK>

F:828-836/Region: protein kinase ATP-binding motif

Query Match 62.9%; Score 2432.5; DB 2; Length 1123;  
Best Local Similarity 68.1%; Pred. No. 9.6e-129;  
Matches 441; Conservative 50; Mismatches 94; Indels 81; Gaps 14.

```

QY 1 MDSIASIVLCGVSLISGVTEGAMDLILNSLPLVSDAETSLTICIASGWPHPEPTIGRD 60
DB 1 MDSLAGLVLCGVSLLYGVVEGAMDLILNSLPLVSDAETSLTICIASGWPHPEPTIGRD 60
QY 61 FEALMNQHQDPLEVTQVTFPEWAKVVKPEKASKINAYFCGSRVPGEAIPRTMKMPQ 120
DB 61 FEALMNQHQDPLEVTQVTFPEWAKVVKPEKASKINAYFCGSRVPGEAIPRTMKMPQ 120
QY 121 QASFLPATLMTVDKGNVNSIFKKVLIKEEDAVIYKNGSFIHSVPHEVDPDILEVILPH 180
DB 121 QASFLPATLMTVDKGNVNSIFKKVLIKEEDAVIYKNGSFIHSVPHEVDPDILEVILPH 180
QY 181 AQPDAGVYSARYIGGNLFTSAPTRILVRKCEAKWCPGPNHLCACMNNGVCHEDTGE 240
DB 181 AQPDAGVYSARYIGGNLFTSAPTRILVRKCEAKWCPGPNHLCACMNNGVCHEDTGE 240
QY 241 ICPEGEMKICEKACELHTEFRICKERKESQEGKSKVPELPPPYGSSAPKWKGLQANE 300
DB 241 ICPEGEMKICEKACELHTEFRICKERKESQEGKSKVPELPPPYGSSAPKWKGLQANE 300
QY 301 AHPHFVCEPKLEPSNNSEMDPEFACGLISPAWGLQETEREILPHATPKIVGLFGLIE 360
DB 301 AHPHFVCEPKLEPSNNSEMDPEFACGLISPAWGLQETEREILPHATPKIVGLFGLIE 360
QY 361 VNSCKFNPTCKASGWLPTNEMILVLPDGLVLPKDFNHTDQFSAVLETHLILDTSC 420
DB 361 VNSCKFNPTCKASGWLPTNEMILVLPDGLVLPKDFNHTDQFSAVLETHLILDTSC 420
QY 421 VVVSVNTVAGWKEKPNISVKKVLPKPLNPNVITDGNFAVINISSEFYFS----- 472
DB 421 VVGSVNTVAGWKEKPNISVKKVLPKPLNPNVITDGNFAVINISSEFYFS----- 472
QY 473 -----EPKS-----CDKTHICPPCPAPELLGSP-----S 496
DB 473 -----EPKS-----CDKTHICPPCPAPELLGSP-----S 496
QY 497 VHLFPPKPKDLMISRTPLFVTCVVVDVSHEDPEVKFNWYVDGVGVHNAKTKPREQYNS 555
DB 497 VHLFPPKPKDLMISRTPLFVTCVVVDVSHEDPEVKFNWYVDGVGVHNAKTKPREQYNS 555
QY 539 IGLPPLPGLSLPKSQTALNTQPIETNSD-----EFY----VEVERPSLQTTSDQN- 589
DB 539 IGLPPLPGLSLPKSQTALNTQPIETNSD-----EFY----VEVERPSLQTTSDQN- 589
QY 556 TYRVSVLT-VLHGDWLNKREYKVKSNKALPAVLETKISKAKQV-PRFQVVT----LP 609
DB 556 TYRVSVLT-VLHGDWLNKREYKVKSNKALPAVLETKISKAKQV-PRFQVVT----LP 609
QY 590 -IKVFGNLTSLVLSNIVTRFGVTVRA-----RVNTKAQGRWSEMLRAWTLSLP 638
DB 590 -IKVFGNLTSLVLSNIVTRFGVTVRA-----RVNTKAQGRWSEMLRAWTLSLP 638
QY 610 PSREEM-----TKNVSSTLCVKGFPSPDIAVWESNCGPENNY 648
DB 610 PSREEM-----TKNVSSTLCVKGFPSPDIAVWESNCGPENNY 648
QY 639 PQPENIKSLNIDTAMVSWT-IVDGYSSISIIIRYKVGKQKNEQD 684
DB 639 PQPENIKSLNIDTAMVSWT-IVDGYSSISIIIRYKVGKQKNEQD 684

```

RESULT 4

JN0712

N:protein-tyrosine kinase (EC 2.7.1.112), receptor type I, precursor mouse

C:Species: Mus musculus (house mouse)

C:Date: 27-Jun-1994 #sequence\_revision 27 Jun 1994 #text\_change 16 Jul-1999

C:Accession: JN0712; S57848, S43495, S43494, S33142  
R:Watanabe, A.; Hamaguchi, T.; Hashiyama, M.; Murayama, Y.; Yasunaga, K.; Suda, T.  
Biochem. Biophys. Res. Commun. 195, 301-309, 1993  
A:Title: Molecular cloning and characterization of mouse tie-2 and IER receptor tyrosin  
A:Reference number: JN0712; M015:93371421  
A:Accession: JN0712  
A:Molecule type: mRNA  
A:Residues: 1-1123 <IMA>  
F:Satoh, T.N.; Qiu, Y.; Kozak, C.A.; Audus, K.L.  
Proc. Natl. Acad. Sci. U.S.A. 90, 9355-9358, 1993  
A:Title: tie-2 and Lie-2 define another class of putative receptor tyrosine kinase ge  
A:Reference number: S57845; M015:94023374  
A:Accession: S57848  
A:Status: preliminary, nucleic acid sequence not shown, translation not shown  
A:Molecule type: mRNA  
A:Residues: 1-766 788-1123 <SAT>  
A:Cross-references: EMBL X71426, NID:4296612, EMBL:CAAC557.1, FID:4296613  
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, March 1993  
R:Kumant, D.J.  
submitted to the EMBL Data Library, April 1993  
A:Reference number: S43495  
A:Accession: S43495  
A:Molecule type: mRNA  
A:Residues: 1-537 539-735 737-786 788-1123 <DUM>  
A:Cross-references: EMBL X71426, NID:4296612, EMBL:CAAC557.1, FID:4296613  
R:Kumant, D.J.; Yamaguchi, T.P.; Conlon, R.A.; Rossant, J.; Breitman, M.L.  
Oncogene 7, 1471-1480, 1992  
A:Title: tek, a novel tyrosine kinase gene located on mouse chromosome 4, is expresse  
A:Reference number: S43494; M015:92334855  
A:Accession: S43494  
A:Molecule type: mRNA  
A:Residues: 823-1123 <DUW>  
A:Cross-references: EMBL X57553  
C:Comment: Receptor tyrosine kinase-ligand systems play an important role in the cons

C:Genetics:

A:Map position: 4

C:Superfamily: protein-tyrosine kinase, receptor type III, EGF homology, fibronectin  
C:Keywords: ATP, autophosphorylation, duplication, glycoprotein, phosphoprotein, phos

F:1-25/Domain: signal sequence #status predicted <SIG>

F:23-1123/Region: protein tyrosine kinase, receptor type tie-2 #status predicted <MA

F:37-104/Domain: immunoglobulin homology <IM1>

F:135-137/Region: cell attachment (P-G-D) motif

F:211-251/Domain: EGF homology <EG1>

F:255-298/Domain: EGF homology <EG2>

F:302-340/Domain: EGF homology <EG3>

F:364-426/Domain: immunoglobulin homology <IM2>

F:447-525/Domain: fibronectin type III repeat homology <FN3A>

F:540-624/Domain: fibronectin type III repeat homology <FN3B>

F:637-719/Domain: fibronectin type III repeat homology <FN3C>

F:751-773/Domain: transmembrane #status predicted <TM>

F:821-1098/Domain: protein kinase homology <PK>

F:829-837/Region: protein kinase ATP-binding motif

F:140-158 390,438,444,558,595,648,690/Binding site: carbohydrate (Asn) (covalent) #st

F:854,871,963/Active site: Lys, Glu, Asp #status predicted

Query Match 62.9%; Score 2432.5; DB 1; Length 1123;

Best Local Similarity 68.1%; Pred. No. 9.6e-129;

Matches 441; Conservative 50; Mismatches 94; Indels 81; Gaps 14;

```

QY 1 MDSIASIVLCGVSLISGVTEGAMDLILNSLPLVSDAETSLTICIASGWPHPEPTIGRD 60
DB 1 MDSLAGLVLCGVSLLYGVVEGAMDLILNSLPLVSDAETSLTICIASGWPHPEPTIGRD 60
QY 61 FEALMNQHQDPLEVTQVTFPEWAKVVKPEKASKINAYFCGSRVPGEAIPRTMKMPQ 120
DB 61 FEALMNQHQDPLEVTQVTFPEWAKVVKPEKASKINAYFCGSRVPGEAIPRTMKMPQ 120
QY 121 QASFLPATLMTVDKGNVNSIFKKVLIKEEDAVIYKNGSFIHSVPHEVDPDILEVILPH 180
DB 121 QASFLPATLMTVDKGNVNSIFKKVLIKEEDAVIYKNGSFIHSVPHEVDPDILEVILPH 180
QY 181 AQPDAGVYSARYIGGNLFTSAPTRILVRKCEAKWCPGPNHLCACMNNGVCHEDTGE 240

```

```

Db 181 AOPQAGVYSAFVIGNLFTSAFTPLVPPPEAKWPPSPPTTNNVWHEIVTER 240
QY 241 ICFFGGMCHTCKAKREHIGGTCGRSGGPGCKSYVFLGQVYQSTAGWKGIGONE 300
Db 241 ICFFGGMCHTCKAKREHIGGTCGRSGGPGCKSYVFLGQVYQSTAGWKGIGONE 300
QY 301 ACBPAYGPGCKRCKSCNNKSWDRFGGTCSPGWKGLQCPEKTPKTPKIVLPDIE 360
Db 301 ACBPAYGPGCKRCKSCNNKSWDRFGGTCSPGWKGLQCPEKTPKTPKIVLPDIE 360
QY 461 VNSKKNPKCKASWPLPTNEMTLVKPGTVVTHFFPKNHTHPSVAFTTHPLLPDSG 420
Db 461 VNSKKNPKCKASWPLPTNEMTLVKPGTVVTHFFPKNHTHPSVAFTTHPLLPDSG 420
QY 472 VVWCVSVNTVAGVVEKPFENISVAVLKKLNAYVNDLGNFAVINLSHPTG----- 472
Db 472 VVWCVSVNTVAGVVEKPFENISVAVLKKLNAYVNDLGNFAVINLSHPTG----- 472
QY 480 VVWCVSVNTVAGVVEKPFENISVAVLKKLNAYVNDLGNFAVINLSHPTG----- 480
Db 480 VVWCVSVNTVAGVVEKPFENISVAVLKKLNAYVNDLGNFAVINLSHPTG----- 480
QY 496 -----EPKS-----CDKTHCPGPCAPPELLGGP-----S 496
Db 496 -----EPKS-----CDKTHCPGPCAPPELLGGP-----S 496
QY 538 LFYKPVNQAWKYEVTVNEIFTLNLYLEPRTTYEL---VQLAPPSPGRSGIPGVRRETT 538
Db 538 LFYKPVNQAWKYEVTVNEIFTLNLYLEPRTTYEL---VQLAPPSPGRSGIPGVRRETT 538
QY 555 VFLPEPKPKDTLMISKTP-EVTCVVVDVSDVSDVSDVSDVSDVSDVSDVSDVSDVSD 555
Db 555 VFLPEPKPKDTLMISKTP-EVTCVVVDVSDVSDVSDVSDVSDVSDVSDVSDVSDVSD 555
QY 589 IGLPPPGGLSLPKSOTALNLWQPIFTNSD-----EFY---VEVERRSLOTTSDQO- 589
Db 589 IGLPPPGGLSLPKSOTALNLWQPIFTNSD-----EFY---VEVERRSLOTTSDQO- 589
QY 609 TYKVVSVLT-VLHQWLNKKEKCKVSNKALPAPTEKTISKAKGQ-PRPQVWT----- 609
Db 609 TYKVVSVLT-VLHQWLNKKEKCKVSNKALPAPTEKTISKAKGQ-PRPQVWT----- 609
QY 638 IKVPGNLISVLLSVLPPEQYTVRA-----RVNTRKAGQENSEELPAWLTSLILP 638
Db 638 IKVPGNLISVLLSVLPPEQYTVRA-----RVNTRKAGQENSEELPAWLTSLILP 638
QY 648 PPSREEM-----TKNQVSLTCLVKGFYPSDIAVEWESNGQPENNY 648
Db 648 PPSREEM-----TKNQVSLTCLVKGFYPSDIAVEWESNGQPENNY 648
QY 683 POPENIKISNITDSAMVSWT-IVDGYSSISSIIIRYKVGCKNEQDH 683
Db 683 POPENIKISNITDSAMVSWT-IVDGYSSISSIIIRYKVGCKNEQDH 683

RESULT 5
JH0771
protein-tyrosine kinase (P2.7.1.112), receptor type tyk precursor - mouse
C:Species: Mus musculus (house mouse)
C:Date: 10-Jun-1993 #sequence_revision 10-Jun-1993 #text_change 21-Jul-2000
C:Accession: JH0771
R:Horita, K.; Yagi, T.; Kohmura, N.; Tomooka, Y.; Ikawa, Y.; Aizawa, S.
Biochem. Biophys. Res. Commun. 189, 1747-1753, 1992
A:Title: A novel tyrosine kinase, tyk, expressed in murine embryonic stem cells.
A:Reference number: JH0771; MUID:94129253
A:Gene: tyk
A:Molecule type: mRNA
A:Residues: 1-1125 <HOR>
A:Cross-references: CB 113746 NID 4230439, FIDM 0AA03883 1; FID 4230440
C:Genetics:
A:Gene: tyk
A:Superfamily: protein tyrosine kinase, receptor type tie, EGF homology; fibronectin tyk
C:Keywords: ATP; autophosphorylation; duplication; glycoprotein; phosphoprotein; phospho
F:1-22/Domain: signal sequence #status predicted <SIG>
F:23-115/Domain: protein tyrosine kinase, receptor type tyk #status predicted <MAT>
F:37-104/Domain: immunoglobulin homology <IM>
F:115-137/Domain: cell attachment (R-G-D) motif
F:212-252/Domain: EGF homology <EGF>
F:256-296/Domain: EGF homology <EGF>
F:303-341/Domain: EGF homology <EGF>
F:345-427/Domain: immunoglobulin homology <IM>
F:448-526/Domain: fibronectin type III repeat homology <FN3A>
F:541-625/Domain: fibronectin type III repeat homology <FN3B>
F:638-720/Domain: fibronectin type III repeat homology <FN3C>
F:752-777/Domain: transmembrane #status predicted <TM>
F:823-1100/Domain: protein kinase homology <KIN>
F:831-839/Region: protein kinase Asp-binding motif
F:1140-1584/Region: protein kinase Asp-binding motif
F:856-873/Active site: lys, glu, asp #status predicted

```

Query Match 61.1% Score 2363 DB 1: Length 1125

```

Best Local Similarity 66.8% Pred. No. 7,5e-125
Matches 473; Conservative 52; Mismatches 99; Indels 84; Gaps 16;
QY 1 MSLASLVGCVSLLSGIVGAMDLILLINSPLVSDAELSLICIASWRPHPTTGRD 60
Db 1 MSLASLVGCVSLLSGIVGAMDLILLINSPLVSDAELSLICIASWRPHPTTGRD 60
QY 61 FFALMNHGCTPLEVTSVTFEWAQKVVVWPKFASKINAVFCEPVP-EALIPETMPMPQ 120
Db 61 FFALMNHGCTPLEVTSVTFEWAQKVVVWPKFASKINAVFCEPVP-EALIPETMPMPQ 120
QY 121 QASFLPATLMTVDKGDNNVNSFKKVLKEEDAVIYKNGSPH--SVPRHEVDPDLEVHL 178
Db 121 QASFLPATLMTVDKGDNNVNSFKKVLKEEDAVIYKNGSPH--SVPRHEVDPDLEVHL 178
QY 179 PHAQVQDAGVSARYLQGNLFTSAFTPLVPPPEAKWPPSPPTTNNVWHEIVTER 238
Db 179 PHAQVQDAGVSARYLQGNLFTSAFTPLVPPPEAKWPPSPPTTNNVWHEIVTER 238
QY 239 EPTPEPPEKMPRTCKAKCELHTFTGPTCKEPAQSGQKSVFLDPEYQSCATGKRLQG 298
Db 239 EPTPEPPEKMPRTCKAKCELHTFTGPTCKEPAQSGQKSVFLDPEYQSCATGKRLQG 298
QY 299 NEAHTPTVTFCTYTP-SNN-EMTRFPGTTPSP-WCH-LPEDEGTEPMTKTVGLLEFH 358
Db 299 NEAHTPTVTFCTYTP-SNN-EMTRFPGTTPSP-WCH-LPEDEGTEPMTKTVGLLEFH 358
QY 359 IFVNSCKNPCKASWPLPTNEMTLVKPGTVVTHFFPKNHTHPSVAFTTHPLLPD 418
Db 359 IFVNSCKNPCKASWPLPTNEMTLVKPGTVVTHFFPKNHTHPSVAFTTHPLLPD 418
QY 419 SGVWCVSVNTVAGVVEKPFENISVAVLKKLNAYVNDLGNFAVINLSHPTG----- 472
Db 419 SGVWCVSVNTVAGVVEKPFENISVAVLKKLNAYVNDLGNFAVINLSHPTG----- 472
QY 473 -----EPKS-----CDKTHCPGPCAPPELLGGP----- 495
Db 473 -----EPKS-----CDKTHCPGPCAPPELLGGP----- 495
QY 496 LFYKPVNQAWKYEVTVNEIFTLNLYLEPRTTYEL---VQLAPPSPGRSGIPGVRRETT 537
Db 496 LFYKPVNQAWKYEVTVNEIFTLNLYLEPRTTYEL---VQLAPPSPGRSGIPGVRRETT 537
QY 538 ASIGLPPPGGLSLPKSOTALNLWQPIFTNSD-----EFY---VEVERRSLOTTSDQO 589
Db 538 ASIGLPPPGGLSLPKSOTALNLWQPIFTNSD-----EFY---VEVERRSLOTTSDQO 589
QY 589 N-IRKVPGNLISVLLSVLPPEQYTVRA-----RVNTRKAGQENSEELPAWLTSLILP 637
Db 589 N-IRKVPGNLISVLLSVLPPEQYTVRA-----RVNTRKAGQENSEELPAWLTSLILP 637
QY 638 LPPSREEM-----TKNQVSLTCLVKGFYPSDIAVEWESNGQPENNY 648
Db 638 LPPSREEM-----TKNQVSLTCLVKGFYPSDIAVEWESNGQPENNY 648
QY 683 LFFQFENIKISNITDSAMVSWT-IVDGYSSISSIIIRYKVGCKNEQDH 684
Db 683 LFFQFENIKISNITDSAMVSWT-IVDGYSSISSIIIRYKVGCKNEQDH 684

RESULT 6
S69339
Iq heavy chain V region precursor - human
C:Species: Homo sapiens (man)
C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 91-Dec-2000
C:Accession: S69339; S72664
R:Khamilichi, A.A.; Aucouturier, P.; Preud'homme, J.L.; Coeune, M.
Eur. J. Biochem. 225, 14-23, 1995
A:Title: Structure of abnormal heavy chains in human heavy chain deposition disease.
A:Reference number: S69339; MUID:95262687
A:Accession: S69339
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-374 <KHA>
A:Cross-references: EMBL:X81695
R:Khamilichi, A.A.
A:Reference number: S72664
A:Accession: S72664
A:Status: preliminary
A:Molecule type: mRNA

```





[illegible]

R:Holmann, T.; Parr, D.M.  
Mol. Immunol. 16, 923-925, 1979  
A:Title: A note on the amino acid sequence of residues 381-391 of human immunoglobulin G  
A:Reference number: A9142; MUID:80114415  
A:Contents: Zie  
A:Accession: A9132  
A:Molecule type: protein  
A:Residues: 238-275, 423-424  
R:Holmann, T.; Parr, D.M.  
submitted to the Atlas, March 1980  
A:Reference number: A94591  
A:Contents: annotation; this revision to residues 25, 69, 49, and 264-268  
A:Note: the revised sequence differs from that shown in having 60 Ala and in the amidated  
residues  
R:Milstein, C.; Fraughton, B.  
Biochem. J. 121, 217-225, 1971  
A:Title: Disulphide bridges of the heavy chain of human immunoglobulin G2.  
A:Reference number: A90259; MUID:72020300  
A:Contents: annotation; myeloma protein Sa, disulfide bonds  
R:Frangione, B.; Milstein, C.; Pink, J.R.L.  
Nature 221, 145-146, 1969  
A:Title: Structural studies of immunoglobulin G  
A:Reference number: A93157; MUID:69064124  
A:Contents: annotation; Sa, disulfide bonds  
C:Genetics:  
A:Gene: GMB, IGHC2  
A:Cross-references: GDB:119348; OMIM:147110  
A:Map position: 1432.33 1432.73  
A:Complex: An immunoglobulin heterotetramer subunit consists of two identical light (L) and heavy (H) chains. In some cases, such as IgA and IgM, the subunits associate into higher order structures. In some cases, such as IgA and IgM, the subunits associate into higher order structures. In some cases, such as IgA and IgM, the subunits associate into higher order structures.  
C:Suprafamily: immunoglobulin C region, immunoglobulin homology  
C:Keywords: duplication, glycoprotein, heterotetramer, immunoglobulin  
F:20-85/Domain: immunoglobulin homology <IM2>  
F:134-203/Domain: immunoglobulin homology <IM3>  
F:240-307/Domain: immunoglobulin homology <IM3>  
F:14/Disulfide bonds: interchain (to light chain) #status experimental  
F:27-83,141,201,247,305/Disulfide bonds: #status predicted  
F:102,103,106,109/Disulfide bonds: interchain (to heavy chain) #status experimental  
F:176/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 29.6% Score 1154; DB 1; Length 326;  
Best Local Similarity 82.4% Pred. No. 1e-57;  
Matches 215; Conservative 13; Mismatches 15; Gaps 2.

QY 461 AVINISSEYFSEPKSKIKH - - - - - TCCPCAPPELLGSPVFLPPPK 503  
DB 67 SVTVVSSNFGTQYTCNWDKPSNKKVKTVERKCKQVECPHPAPP-VAGSPVFLPPPK 125  
QY 504 PKDTLMSIRPVTCTVVDVSHEDPEVFNWYVGVVEVHNKIKPREQKNSIYRVVSVL 563  
DB 126 PKDTLMSIRPVTCTVVDVSHEDPEVFNWYVGVVEVHNKIKPREQKNSIYRVVSVL 185  
QY 564 IVHQWLNGKLYKCKVSNNAIPAPTEKTSKAKGQPPQVYTLPPSPPEMTKNQVSLI 623  
DB 186 IVHQWLNGKLYKCKVSNKGLPAPTEKTSKAKGQPPQVYTLPPSPPEMTKNQVSLI 245  
QY 624 CLVKGFPYSDIAVENESKCPENNYKTPFPVLDSDGSEFLYSLKLTVDKSRWQNVISCS 683  
DB 246 CLVKGFPYSDIAVENESKCPENNYKTPFPVLDSDGSEFLYSLKLTVDKSRWQNVISCS 705  
QY 684 VMHEALHNHYTKSLSLSPGK 704  
DB 306 VMHEALHNHYTKSLSLSPGK 326

RESULT 14  
GHRH  
Ig gamma 4 chain C region - human  
C:Species: Homo sapiens (man)  
C:Date: 02-Apr-1992 #sequence revision 92 Apr-1982 #text\_change 16-Jul-1999  
C:Accession: A90933; A90249; A92150  
R:Ellison, J.; Huxbaum, J.; Hood, L.

DNA 1, 11-18, 1981  
A:Title: Nucleotide sequence of a human immunoglobulin C gamma4 gene.  
A:Reference number: A90933; MUID:84157104  
A:Accession: A90933  
A:Molecule type: DNA  
A:Residues: 1-327 <ELL>  
A:Note: the sequence was determined from the genome gene  
R:Pink, J.R.L.; Buttery, S.H.; De Vries, G.M.; Milstein, C.  
Biochem. J. 117, 33-47, 1970  
A:Title: Human immunoglobulin subclasses. Partial amino acid sequence of the constant  
A:Reference number: A90249; MUID:70207560  
A:Accession: A90249  
A:Molecule type: protein  
A:Residues: 1-30; 81-326 <PIN>  
C:Genetics:  
A:Gene: GDB, IGHC4  
A:Cross-references: GDB:119340; OMIM:147110  
A:Map position: 1432.33 1432.33  
A:Introns: 99/1; 111/1; 221/1  
C:Complex: An immunoglobulin heterotetramer subunit consists of two identical light (L) and heavy (H) chains. In some cases, such as IgA and IgM, the subunits associate into higher order structures. In some cases, such as IgA and IgM, the subunits associate into higher order structures. In some cases, such as IgA and IgM, the subunits associate into higher order structures.  
C:Suprafamily: immunoglobulin C region, immunoglobulin homology  
C:Keywords: duplication, glycoprotein, heterotetramer, immunoglobulin  
F:20-85/Domain: immunoglobulin homology <IM2>  
F:134-203/Domain: immunoglobulin homology <IM2>  
F:240-307/Domain: immunoglobulin homology <IM3>  
F:14/Disulfide bonds: interchain (to light chain) #status experimental  
F:27-83,141,201,247,305/Disulfide bonds: #status predicted  
F:102,103,106,109/Disulfide bonds: interchain (to heavy chain) #status experimental  
F:177/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 29.6% Score 1144; DB 1; Length 327;  
Best Local Similarity 89.9% Pred. No. 3.8e-57;  
Matches 213; Conservative 7; Mismatches 9; Gaps 1;

QY 468 HYPCEPKSKIKHICPCAPPELLGSPVFLPPPKKIDLMISRPPEVTCVVDVSHED 527  
DB 99 FSKYGP - - - - - CPSCAPPEFLGSPVFLPPPKKIDLMISRPPEVTCVVDVSHED 150  
QY 528 PVKKNWYVGVVEVHNKIKPREQKNSIYRVVSVLIVHQWLNGKLYKCKVSNKGLP 587  
DB 151 PVKKNWYVGVVEVHNKIKPREQKNSIYRVVSVLIVHQWLNGKLYKCKVSNKGLP 210  
QY 588 PTEKTSKAKGQPPQVYTLPPSPPEMTKNQVSLTCLVKGFPYSDIAVENESKCPEN 647  
DB 211 STEKTSKAKGQPPQVYTLPPSPPEMTKNQVSLTCLVKGFPYSDIAVENESKCPEN 270  
QY 648 YKTPPVLDSDGSEFLYSLKLTVDKSRWQNVISCSVMHEALHNHYTKSLSLSPGK 704  
DB 271 YKTPPVLDSDGSEFLYSLKLTVDKSRWQNVISCSVMHEALHNHYTKSLSLSPGK 327

RESULT 15  
GHRH  
Ig gamma chain C region - rabbit  
C:Species: Oryctolagus cuniculus (domestic rabbit)  
C:Date: 24-Apr-1984 #sequence revision 15-Nov-1984 #text\_change 16-Jul-1999  
C:Accession: A91749; A90240; A90248; A90245; A94416; A02161  
R:Bernstein, K.E.; Alexander, C.B.; Mage, R.G.  
Immunogenetics 18, 387-397, 1983  
A:Title: Nucleotide sequence of a rabbit IgG heavy chain from the recombinant F1 hap  
A:Reference number: A91749; MUID:84030930  
A:Accession: A91749  
A:Molecule type: mRNA  
A:Residues: 1-323 <BER>  
A:Note: this sequence has the 412 allelic marker 104 (Thr) and the 414 marker, 145-  
R:Pratt, D.M.; Mole, L.E.  
Biochem. J. 151, 337-349, 1975  
A:Title: Sequence studies on the constant region of the Fd sections of rabbit immunog  
A:Reference number: A90290; MUID:76135469  
A:Accession: A90290

A:Molecule type: protein  
 A:Residues: 1-47, 'E', 49-71, 'FV', 72-128, 'FEA'  
 R:Martens, C.L.; Moore, K.W.; Steinmetz, M.; Hood, L.; Knight, K.L.  
 Proc. Natl. Acad. Sci. U.S.A. 79: 6018-6022, 1982  
 A:Title: Heavy chain genes of rabbit IgG: Isolation of a cDNA encoding gamma heavy chain  
 A:Reference number: A93928; MUID:83299917  
 A:Accession: A93928  
 A:Molecule type: mRNA  
 A:Residues: 48-103, 'M', 105-143, 'P', 145-184, 'A', 186, 'E', 198-256, 'MAR'  
 A:Cross-references: CHM16426; NID-0165111; PIRN:AAA128921; PIR:0165112  
 A:Note: this sequence has the d11 allotypic marker, 104-Met, and the e15 allotypic marker  
 R:Euchner, R.G.; Jackson, S.A.; Mole, L.E.; Porter, P.P.  
 Biochem. J. 116: 249-259, 1970  
 A:Title: Sequence studies of the Fd section of the heavy chain of rabbit immunoglobulin  
 A:Reference number: A90245; MUID:70110015  
 A:Accession: A90245  
 A:Molecule type: protein  
 A:Residues: 132-143, 'E', 145-161, 'PRD'  
 R:Hull, R.L.; Lebovitz, H.E.; Fellows Jr., R.E.; Delaney, R.  
 in Gamma Globulins, N. L. S. A. Ellstrand, J. L. Ed., pp. 109-127, Almqvist and Wiksell,  
 A:Reference number: A94416  
 A:Accession: A94416  
 A:Molecule type: protein  
 A:Residues: 129-131, 155-172, 'P', 174-184, 'A', 186, 'E', 188-209, 'D', 202-212, 'E', 219-232, 'Q'  
 A:Note: this has the e15 allotypic marker, 185-Ala  
 C:Complex. An immunoglobulin heterotetramer subunit consists of two identical light (kappa) and two identical heavy (lambda) chains. In some cases, such as IgA and IgM, the subunits associate into higher order structures.  
 C:Superfamily. Immunoglobulin C region; immunoglobulin homology  
 C:Keywords. duplication, glycoprotein, heterotetramer; immunoglobulin  
 F:26-82/Domains: immunoglobulin homology <IM1>  
 F:130-169/Domains: immunoglobulin homology <IM2>  
 F:235-303/Domains: immunoglobulin homology <IM3>  
 F:173/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 24.1%; Score 932; DB 1; Length 323;  
 Best Local Similarity 59.5%; Pred. No. 2.8e-45;  
 Matches 173; Conservative 40; Mismatches 48; Indels 30; Gaps 3;  
 QY 444 LPKPLNA-----PNI DTGHNPAVINISSEPFYGFGRKSKDTH-----481  
 DB 33 LPEPVTITWNSGTLTNGVTFPSVPQSSGLYSLSSVSVTSQQVPTCNVHIPATNTKVD 92  
 QY 482 -----TC--PPCPAPELLGGSPVFLPPPKKDTLMSIRTPETCVVVDVSHEDPEVKEN 533  
 DB 93 KTVAPSTCSKPTCTPPHLLGGSPVFLPPPKKDTLMSIRTPETCVVVDVSDQDDPEVQT 152  
 QY 534 WYDGVVEVHNAKTPREQYNSTYRVVSVLTVLHODWLNKGYKCKVSKALPAPIEKTI 593  
 DB 153 WYINNEQVRTAPPPIPEQFNSTIPVVSITLPIHODWLNKGYKCKVSKALPAPIEKTI 212  
 QY 594 SKAGQPREPOVYIIPSRREMKNQVSLICLVKGYPSDIAVWFNSGQPENNYKTPPP 653  
 DB 213 SKAPGPLEKVVIMPPPEELSPSPSLTMINGEYPSDISVVEWFKNAEENYKTTFA 272  
 QY 654 VLSDGSGFELYSLTVKDSKQGNVPSQVSHALHNHYTKSLSPGK 704  
 DB 273 VLSDGSGFELYSLTVKDSKQGNVPSQVSHALHNHYTKSLSPGK 323

Search completed: September 23, 2002, 22:11:26  
 Job time: 162 sec











RT The endothelial-specific receptor tyrosine kinase, tek, is a member  
 RT of a new subfamily of receptors.\*;  
 RL Oncogene 8:1293-1301(1993).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-Embryonic stem cells;  
 RX MEDLINE-93129253; PubMed-1242811;  
 RA Horita K., Yagi T., Kohmura N., Tomooka Y., Ikawa Y., Alzawa S.;  
 RT "A novel tyrosine kinase, htk, expressed in murine embryonic stem  
 RT cells.\*;  
 RL Biochem. Biophys. Res. Commun. 189:1747-1753(1992).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-Lung;  
 RX MEDLINE-94031116; PubMed-8217221;  
 RA Kunking A.S., Stacker S.A., Wilks A.F.;  
 RT "Tie2, a putative protein tyrosine kinase from a new class of cell  
 RT surface receptor.\*;  
 RL Growth Factors 9:99-105(1993).  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE-94244487; PubMed-8187650;  
 RA Schmeichel H., Risau W.;  
 RT "Expression of tie-2, a member of a novel family of receptor tyrosine  
 RT kinases, in the endothelial cell lineage.\*;  
 RL Development 119:957-968(1993).  
 RN [6]  
 RP SEQUENCE OF 822-1122 FROM N.A.  
 RC STRAIN-CD-1; TISSUE-embryonic heart;  
 RX MEDLINE-92434855; PubMed-1630810;  
 RA Dumont D.J., Yamauchi T.P., Coulon R.A., Rossant J., Breilman M.L.;  
 RT "tek, a novel tyrosine kinase gene located on mouse chromosome 4, is  
 RL expressed in endothelial cells and their presumptive precursors.\*;  
 RL Oncogene 7:1471-1480(1992).  
 CC [1] FUNCTION: THIS PROTEIN IS A PROTEIN TYROSINE KINASE TRANSMEMBRANE  
 CC RECEPTOR FOR ANGIOPOIETIN 1. IT MAY CONSTITUTE THE EARLIEST  
 CC MAMMALIAN ENDOTHelial CELL LINEAGE MARKER. PROBABLY REGULATES  
 CC ENDOTHelial CELL PROLIFERATION, DIFFERENTIATION AND GUIDES THE  
 CC PROPER PATTERNING OF ENDOTHelial CELLS DURING BLOOD VESSEL  
 CC FORMATION.  
 CC [2] CATALYTIC ACTIVITY: ATP + a protein tyrosine -> ADP + protein  
 CC tyrosine phosphate.  
 CC [3] SUBCELLULAR LOCATION: Type I membrane protein  
 CC [4] TISSUE SPECIFICITY: SPECIFICALLY EXPRESSED IN DEVELOPING VASCULAR  
 CC ENDOTHelial CELLS.  
 CC [5] DEVELOPMENTAL STAGE: EXPRESSION DETECTABLE IN DAY 8.5 EMBRYOS.  
 CC [6] SIMILARITY: TO OTHER PROTEIN TYROSINE KINASES IN THE CATALYTIC  
 CC DOMAIN.  
 CC [7] SIMILARITY: CONTAINS 2 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.  
 CC [8] SIMILARITY: CONTAINS 3 FIBRONECTIN TYPE III-LIKE DOMAINS.  
 CC [9] SIMILARITY: CONTAINS 3 EGF-LIKE DOMAINS.  
 CC [10] This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement. (See <http://www.isb.sib.ch/announcement/>  
 CC or send an email to [license@sib.sib.ch](mailto:license@sib.sib.ch).)  
 CC [11] EMBL: X71426; CAA50557.1;  
 CC EMBL: X67553; CAA47857.1;  
 CC EMBL: D13738; BAA02883.1;  
 CC EMBL: S67051; AAB28663.1;  
 CC EMBL: S33142; S33142;  
 CC EMBL: M61; 98664; Tek.  
 CC InterPro: IPR000561; FGF-like  
 CC InterPro: IPR000719; FGF-like  
 CC InterPro: IPR003961; FN III.  
 CC InterPro: IPR001245; Tyr\_kinase.  
 CC Pfam: PF00041; fn3; 3.  
 CC Pfam: PF00069; pkinase; 1.

DP PRINTS: P000109; TYRKINASE.  
 DR SMART: SM00181; EGF\_2.  
 DR SMART: SM00640; FN3\_2.  
 DR SMART: SM00219; TYRK\_1.  
 DR PROSITE: PS00107; PROTEIN\_KINASE\_ATP; 1.  
 DR PROSITE: PS00109; PROTEIN\_KINASE\_TYR; 1.  
 DR PROSITE: PS00611; PROTEIN\_KINASE\_DOM; 1.  
 DR PROSITE: PS00622; EGF\_1; 3.  
 DR PROSITE: PS01186; EGF\_2; 3.  
 KW Receptor, tyrosine protein kinase, transferase; Signal: AIP-binding;  
 KW Repeat, EGF-like domain; Transmembrane; Immunoglobulin domain;  
 KW Glycoprotein; Phosphorylation; Multigene family.  
 FT SIGNAL 1 18  
 FT CHAIN 19 1122  
 FT DOMAIN 19 744  
 FT TRANSMEM 745 769  
 FT DOMAIN 770 1122  
 FT DOMAIN 44 102  
 FT DOMAIN 210 252  
 FT DOMAIN 254 299  
 FT DOMAIN 301 341  
 FT DOMAIN 370 424  
 FT DOMAIN 444 536  
 FT DOMAIN 539 633  
 FT DOMAIN 637 731  
 FT DOMAIN 822 1094  
 FT NP\_HIND 828 836  
 FT ACT\_SITE 853 853  
 FT ACT\_SITE 962 962  
 FT CARBOHYD 140 140  
 FT CARBOHYD 158 158  
 FT CARBOHYD 399 399  
 FT CARBOHYD 438 438  
 FT CARBOHYD 464 464  
 FT CARBOHYD 558 558  
 FT CARBOHYD 595 595  
 FT CARBOHYD 648 648  
 FT CARBOHYD 690 690  
 FT CARBOHYD 990 990  
 FT CONFLICT 161 171  
 FT CONFLICT 538 538  
 FT CONFLICT 736 736  
 FT CONFLICT 745 761  
 FT CONFLICT 786 786  
 FT CONFLICT 913 913  
 FT CONFLICT 925 931  
 FT CONFLICT 1117 1117  
 SQ SEQUENCE 1122 AA; 125700 MW; F879640103FE96 CRC64;  
 Query Match 62.9%, Score 2432.5; DH 1; LCsth; 1122;  
 Best Local Similarity 68.1%; Pred No 1 60-161;  
 Matches 481; Conservative 50; Mismatches 94; Indels 81; Gaps 14;  
 QY 1 MESLASLVGVSVLLSGVSGAMULLINSLPVSAGAEISLGLASGWRPHEPTLIGED 60  
 DB 1 MESLASLVGVSVLLSGVSGAMULLINSLPVSAGAEISLGLASGWRPHEPTLIGED 60  
 QY 61 FEALMNHQDPLEVTQVTPREWKKVVKPKASKINFAVFCEGVPSEAIPIPTWPMQ 120  
 DB 61 FEALMNHQDPLEVTQVTPREWKKVVKPKASKINFAVFCEGVPSEAIPIPTWPMQ 120  
 QY 121 QASFLPATLTMTVDKGVNVI SPKKVLKEDAVIYKNGSFTHSVPRHVVDFLEVLPH 180  
 DB 121 QASFLPATLTMTVDKGVNVI SPKKVLKEDAVIYKNGSFTHSVPRHVVDFLEVLPH 180  
 QY 181 AGQPAWVSAPYIGGNI FTSAFTPLIV9PCEAGKFWGPECNHLCTACNNVAVVHEITGEC 240  
 DB 181 AGQPAWVSAPYIGGNI FTSAFTPLIV9PCEAGKFWGPECNHLCTACNNVAVVHEITGEC 240  
 QY 241 TCPPGFMGRTCEKACELHFTGRTCKERSGEGKSVVPCLPDPYGGSTATGKGLQCN 400  
 DB 241 TCPPGFMGRTCEKACELHFTGRTCKERSGEGKSVVPCLPDPYGGSTATGKGLQCN 400



```

FI DISULFID 109 109 INTERCHAIN (WITH HEAVY CHAIN).
FI DISULFID 112 112 INTERCHAIN (WITH HEAVY CHAIN).
FI DISULFID 144 204
FI DISULFID 250 308
FI CARBOHYD 140 180
FI MOD_RES 330 330
FI VARIANT 97 97
FI VARIANT 259 259
FI VARIANT 241 241
FI STRAND 123 126
FI HELIX 130 134
FI TURN 146 137
FI STRAND 141 148
FI STRAND 158 162
FI TURN 163 164
FI STRAND 165 165
FI STRAND 175 178
FI STRAND 184 190
FI HELIX 193 197
FI TURN 198 199
FI STRAND 202 206
FI STRAND 215 219
FI STRAND 227 227
FI STRAND 230 234
FI HELIX 248 240
FI TURN 241 242
FI STRAND 245 256
FI STRAND 260 266
FI TURN 267 268
FI STRAND 259 270
FI STRAND 274 276
FI STRAND 280 281
FI TURN 283 284
FI STRAND 287 296
FI HELIX 297 301
FI TURN 302 303
FI STRAND 306 312
FI TURN 313 314
FI TURN 316 317
FI STRAND 320 324
SQ SEQUENCE 330 AA 36106 MW 3770PF106C2P4A30 C9C64.

Query Match 32.6%; Score 1262; DB 1; Length 330;
Best local Similarity 78.8%; Pred. No. 8,9c 81;
Matches 242; Conservative 12; Mismatches 11; Indels 42; Gaps 5.

QY 405 SVAITTHIRLP FDSGV WVCSTVTAGMVERKFNISV KVLKPLNAPNVIDTG 457
Db 59 SSGLYSLSSWTVPSSSLSGTUTYICNVN-----HKPSNTRKDKKV----- 98
QY 458 HNFVAVINISSEHYFCEPKSCDKIHCHPCAPPELLGGPSVFLPPPKKDILMISRIPEVT 517
Db 99 -----EPKSCDKIHCHPCAPPELLGGPSVFLPPPKKDILMISRIPEVT 143
QY 518 CVVDVDSHEDPEVKFNWYVGVVEVINAKTRPPFEYNSVTPVWSVLTVLHGMNLSKEYK 577
Db 144 CVVDVDSHEDPEVKFNWYVGVVEVINAKTRPPFEYNSVTPVWSVLTVLHGMNLSKEYK 203
QY 578 QVSNKALPAPTEKTSKAKGTPPEPQVYTLPPSPPEMTRKQVSLTGLVKGFPESDIIVE 637
Db 204 QVSNKALPAPTEKTSKAKGTPPEPQVYTLPPSPPEMTRKQVSLTGLVKGFPESDIIVE 263
QY 638 WESNGQPLNNYKTPVPLDSDGSEFLYSKLVIVKSRWQCNQVTSVCSVHREALNNHYIQKS 697
Db 264 WESNGQPLNNYKTPVPLDSDGSEFLYSKLVIVKSRWQCNQVTSVCSVHREALNNHYIQKS 323
QY 698 LSLSPGR 704
Db 324 LSLSPGR 330

```

```

RESULT 5
GC3_HUMAN
ID GC3_HUMAN STANDARD: PRT: 290 AA.
AC P01860;
DT 21-JUL-1986 (Rel. 01, Created)
DI 21-JUL-1986 (Rel. 01, Last sequence update)
DI 16-OCT-2001 (Rel. 40, Last annotation update)
DE Ig gamma-3 chain C region (heavy chain disease protein) (HHC)
GN IGHG3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE (DISEASE PROTEIN WIS).
RX MEDLINE=81021548; PubMed=6774747;
RA Frangione B., Rosenwasser E., Prelli F., Franklin E.C.;
RT "Primary structure of human gamma 3 immunoglobulin deletion mutant;
RT gamma 3 heavy-chain disease protein Wis.";
RL Biochemistry 19:4304-4308(1980).
RN [2]
RP REVISIONS TO 12-97 OF PROTEIN WIS.
RX MEDLINE=77118561; PubMed=402363;
RA Michaelisen T.E., Frangione B., Franklin E.C.;
RT "Primary structure of the 'hinge' region of human IgA3. Probable
RT quadruplication of a 15-amino acid residue basic unit.";
RL J. Biol. Chem. 252:883-889(1977).
RN [3]
RP REVISIONS TO 59-289 OF PROTEIN WIS (DISEASE PROTEIN ZUC).
RX MEDLINE=77021516; PubMed=823945;
RA Wolfenstein-Todel C., Frangione B., Prelli F., Franklin E.C.;
RT "The amino acid sequence of 'heavy chain disease' protein ZUC.
RT Structure of the Fc fragment of immunoglobulin G3.";
RL Biochem. Biophys. Res. Commun. 71:907-914(1976).
RN [4]
RP SEQUENCE FROM N.A. (DISEASE PROTEIN OMM).
RX MEDLINE=82247835; PubMed=6808505;
RA Alexander A., Steinmetz M., Barrilault D., Frangione B.,
RA Franklin E.C., Hood L., Buxbaum J.N.;
RT "Gamma Heavy chain disease in man: cDNA sequence supports partial
RT gene deletion model.";
RL Proc. Natl. Acad. Sci. U.S.A. 79:3250-3264(1982).
CC 1- SUBUNIT: DIMER LINKED BY 12 DISULFIDE BONDS; IT HAS AN EXTRA
CC INTERCHAIN DISULFIDE BOND AT POSITION 7 IN ADDITION TO THE 11
CC NORMALLY PRESENT IN THE HINGE REGION.
CC 1- MISCELLANEOUS: THE HEAVY CHAIN DISULFIDE PROTEIN WIS IS SHOWN.
CC 1 MISCELLANEOUS: THE SEQUENCE OF RESIDUES 42-76 WAS TAKEN FROM THE
CC REF.2.
CC 1- MISCELLANEOUS: DISEASE PROTEIN WIS IS LACKING MOST OF THE V REGION
CC AND ALL OF THE CH1 REGION.
CC 1- MISCELLANEOUS: DISEASE PROTEIN ZUC LACK MOST OF THE V REGION, ALL
CC OF THE CH1 REGION, AND PART OF THE HINGE COMPARED WITH NORMAL.
CC GAMMA-3 HEAVY CHAINS.
CC 1- MISCELLANEOUS: DISEASE PROTEIN OMM MAY REPRESENT AN ALLELIC FORM
CC OR ANOTHER GAMMA CHAIN SUBCLASS.
CC 1- MISCELLANEOUS: THE HINGE REGION IN GAMMA-3 CHAINS IS ABOUT FOUR
CC LINES AS LONG AS IN OTHER GAMMA CHAINS AND CONTAINS THREE
CC IDENTICAL 15-RESIDUE SEGMENTS PRECEDED BY A SIMILAR 17-RESIDUE
CC SEGMENT (12-28).
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: J09231; AAA52805.1; ALT_SEQ.
DR PIR: A02149; G3H0W1.
DR HSPD: P01857; 1FC1.

```

```

DR MIM: 147120;
DR InterPro: IPR001006; Ig_MHC;
DR InterPro: IPR001597; Ig_L1;
DR InterPro: IPR003600; Ig_Like;
DR Pfam: PF00047; Ig; 2;
DR SMART: SM00410; Ig_Like; 1;
DR SMART: SM00407; IG_L1; 1;
DR PROSITE: PS00290; IG_MHC; 1;
KW Immunoglobulin domain; Immunoglobulin c region; Glycoprotein; Repeat
FT DOMAIN 12 73
FT DOMAIN 74 183
FT DOMAIN 184 289
FT REPEAT 29 43
FT REPEAT 44 58
FT REPEAT 59 73
FT MOD_RES 1 1
FT MOD_RES 6 6
FT CARBOHYD 7 7
FT DISULFID 7 7
FT DISULFID 24 24
FT DISULFID 27 27
FT DISULFID 33 33
FT DISULFID 39 39
FT DISULFID 42 42
FT DISULFID 48 48
FT DISULFID 54 54
FT DISULFID 57 57
FT DISULFID 63 63
FT DISULFID 69 69
FT DISULFID 72 72
FT CARBOHYD 140 140
FT MOD_RES 290 290
FT VARIANT 126 127
FT VARIANT 134 134
FT VARIANT 139 139
FT VARIANT 182 182
FT VARIANT 227 227
FT VARIANT 227 227
FT VARIANT 279 279
FT VARIANT 279 279
SQ SEQUENCE 230 AA, 3233 MW, 56476.457 kDa, pI=6.4,
Query Match 30.1%; Score 1163; DB 1; Length 290;
Best local Similarity 90.3%; Pred. No. 5,9e-74;
Matches 213; Conservative 11; Mismatches 12; Indels 0; Gaps 0;
QY 469 PYGPEPKSGKTHTPPCPPAPPELLGAPSVLFPFKPKDTLMISRTPEVTCVVDVSHEDP 528
DB 1 ||||| || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
55 PROPEKSCDTPPCPPAPPELLGAPSVLFPFKPKDTLMISRTPEVTCVVDVSHEDP 114
QY 529 EVKFNWVGVKVEFINAKTKPEFQYNSTYKVSVLTVLHGLDGLNGKEYCKVSKNKPAP 588
DB 1 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
115 EVQKFWVGVGVVUNAKTKPEFQYNSTYKVSVLTVLHGLDGLNGKEYCKVSKNKPAP 174
QY 589 IETKTSKAGQPREPQVYTLTPSPPEMTKQGVSLTCLVKGYPESDIAYVWFSNQCPENNY 648
DB 1 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
175 IETKTSKAGQPREPQVYTLTPSPPEMTKQGVSLTCLVKGYPESDIAYVWFSNQCPENNY 234
QY 649 KIPEVVLSTGSEFFLYSKILVKKSPWQGVVVFNSVNMWEALHNYTLKSLSLSPGK 704
DB 1 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
235 NITPEMLDSGSPFLYSKILVKKSPWQGVVVFNSVNMWEALHNYTLKSLSLSPGK 290
RESULT 6
GC2_HUMAN
ID GC2_HUMAN PRI: 326 AA.
AC P01859;

```

```

DT 21-JUL-1986 (Ref. 01, Created)
DT 21-JUL-1986 (Ref. 01, Last sequence update)
DE 16-OCT-2001 (Ref. 49, Last annotation update)
DE Ig gamma-2 chain C region.
GN IGK82.
OS Homo sapiens (Human).
OC Eukaryota, Metazoa, Chordata, Vertebrata, Euteleostomi,
OC Mammalia, Eutheria, Primates, Catarrhini, Hominoidea, Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE OF 2-326 FROM N.A.
RX MEDLINE=82197621; PubMed=6804948;
RA Ellison J.W., Hood L.E.;
RT "Linkage and sequence homology of two human immunoglobulin gamma
PL heavy chain constant region genes.";
PL Proc Natl Acad Sci U S A 79:1044-1048(1982).
PN [2]
RP SEQUENCE OF 88-115 FROM N.A.
RX MEDLINE=8401943; PubMed=6811179;
RA Takahashi N., Ueda S., Obara M., Nikaide T., Nakai S., Honjo T.;
RT "Structure of human immunoglobulin gamma genes: implications for
FT evolution of a gene family.";
FT Cell 29:671-679(1982).
RN [3]
RP SEQUENCE OF 99-177 AND 310-326 FROM N.A.
RX MEDLINE=84235902; PubMed=6306676;
RA Frawinkel N., Pabbits T.H.;
RT "Comparison of the hinge coding segments in human immunoglobulin gamma
PL heavy chain genes and the linkage of the gamma 2 and gamma 4 subclass
RN genes.";
RN EMBO J 1:403-407(1982).
RN [4]
RP SEQUENCE OF 1-325 (MYELOMA PROTEIN TIL).
RX MEDLINE 81007873, PubMed=6774012.
RA Wang A.C., Tung E., Fudenberg H.H.;
RT "The primary structure of a human IgG2 heavy chain: genetic,
FT evolutionary, and functional implications.";
FT J. Immunol. 125:1048-1054(1980).
RN [5]
RP SEQUENCE OF 1-85 AND 132-325 (MYELOMA PROTEIN ZIF).
RX MEDLINE=80001357; PubMed=113060;
RA Connell G.F., Parr D.M., Hofmann T.;
RT "The amino acid sequences of the three heavy chain constant region
PL domains of a human IgG2 myeloma protein.";
PL Can. J. Biochem. 57:758-767(1979).
RN [6]
RP SEQUENCE OF 238-275 (ZIF).
RX MEDLINE=80114419; PubMed=118920;
RA Hofmann T., Parr D.M.;
RT "A note of the amino acid sequence of residues 381-391 of human
PL immunoglobulin gamma chains.";
PL Mol. Immunol. 16:923-925(1979).
RN [7]
RP REVISIONS TO 25; 59; 60 AND 264-268 (ZIF).
RA Hofmann T., Parr D.M.;
RT Submitted (MAR-1980) to the PIR data bank.
RN [8]
RP SEQUENCE OF 1-121 (DOT).
RX MEDLINE=95255298; PubMed=7737190;
RA Stoppini M., Rellotti V., Negri A., Merlini G., Garver F., Ferri G.;
RT "Characterization of the two unique human anti-IL-6 monoclonal
PL immunoglobulins.";
PL Eur J Biochem. 224:886-893(1995).
RN [9]
RP DISULFIDE BONDS.
RX MEDLINE=72033500; PubMed=4940472;
RA Milstein C., Fraenkel B.;
RT "Disulphide bridges of the heavy chain of human immunoglobulin G2.";
PL Biochem. J. 121:217-225(1971).
RN [10]
RP DISULFIDE BONDS.

```

```

EX MEDLINE-64064124; PubMed-5782707;
RA Francone B., Milstein C., Pink J.R.L.;
RT "Structural studies of immunoglobulin G.";
RL Nature 221:145-148(1969).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement. (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch.)
CC
DR EMBL: J00240; AAB59393.1;
DR PIR: A02148; G2H0.
DR HSSP: P01857; IFC1.
DR MIM: 147110;
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003597; Ig_C1.
DR InterPro: IPR003600; Ig_Like.
DR Pfam: PF00047; Ig_3.
DR SMART: SM00410; Ig_Like; 1.
DR SMART: SM00407; IgC1; 2.
DR PROSITE: PS00290; Ig_MHC; 2.
KW Immunoglobulin domain; Immunoglobulin C region.
FT NON_TER 1
FT DOMAIN 1 98
FT DOMAIN 99 110
FT DOMAIN 111 219
FT DOMAIN 220 326
FT DISULFID 14 14
FT DISULFID 27 83
FT DISULFID 102 102
FT DISULFID 103 103
FT DISULFID 106 106
FT DISULFID 109 109
FT DISULFID 140 200
FT DISULFID 246 304
FT SITE 155 156
FT MODRES 326 326
FT VARIANT 60 60
FT CONFLICT 109 109
FT SEQUENCE 326 AA; 35484 MW; 8316878567875a0 Cn64;
Query Match 29.8%; Score 1154; DB 1; Length 326;
Best Local Similarity 82.4%; Pred. No. 2.8e-73;
Matches 215; Conservative 14; Mismatches 15; Indels 18; Gaps 2;
QY 461 AVINISSEPPFCEAKSLDNRH-----LPPCHAPPELLGGSVFLPPPK 503
DB 67 SVVIVFSSNPGFQIYTCNDVHKPSNTKVDKIVRKCCVECHPCFAPP-VAGPSVFLPPPK 125
QY 504 PKDTLMISRTPEVTGVVDVSHEDPEVKFNMYVDGSEVHNAKTKPREQYNSTYRVVSVL 563
DB 126 PKDILMISRTPEVTGVVDVSHEDPEVKFNMYVDGSEVHNAKTKPREQYNSTYRVVSVL 185
QY 564 IVLHDMWLNCKEYKCKVSNKALPAPLTKLISKAKGQPPPEVYTLPPSPREMTKNQVSLT 623
DB 186 TVVHDMWLNCKEYKCKVSNKALPAPLTKLISKAKGQPPPEVYTLPPSPREMTKNQVSLT 245
QY 624 CLVKGFFYSIDIAVWESNCQENNYKTIIPVLDSDGSEFFLYSKITVDKSRWQOCNVPSCS 683
DB 246 CLVKGFFYSIDIAVWESNCQENNYKTIIPVLDSDGSEFFLYSKITVDKSRWQOCNVPSCS 305
QY 684 VMIEALHNHYTKSLSLSPK 704
DB 306 VMIEALHNHYTKSLSLSPK 326
RESULT 7
GC4_HUMAN
ID GC4_HUMAN STANDARD; PRT; 327 AA.
AC P01861;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DE 16-OCT-2001 (Rel. 40, Last annotation update)
DE Ig gamma-4 chain C region.
GN IGHS4.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-83157104; PubMed-6299662;
RA Ellison J.W., Ruxbaum J.N., Hood L.F.;
RT "Nucleotide sequence of a human immunoglobulin C gamma 4 gene.";
RL DNA 1:11-18(1981).
RN [2]
RP SEQUENCE OF 1-30 AND 81-326.
RX MEDLINE-70207560; PubMed-4192699;
RA Pink J.R.L., Huttery S.H., de Vries G.M., Milstein C.;
RT "Human immunoglobulin subclasses. Partial amino acid sequence of the
RT constant region of a gamma 4 chain.";
RL Biochem. J. 117:33-47(1970).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement. (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch.)
CC
EMBL: K01316; AAB59394.1; ALT_INIT.
PIR: A02150; G4H0.
HSSP: P01842; 7FAB.
MIM: 147130;
InterPro: IPR003006; Ig_MHC.
InterPro: IPR003597; Ig_C1.
InterPro: IPR003600; Ig_Like.
Pfam: PF00047; Ig_3.
SMART: SM00410; Ig_Like; 1.
SMART: SM00407; IgC1; 2.
PROSITE: PS00290; Ig_MHC; 2.
KW Immunoglobulin domain; Immunoglobulin C region.
FT NON_TER 1
FT DOMAIN 1 98
FT DOMAIN 99 110
FT DOMAIN 111 220
FT DOMAIN 221 327
FT DISULFID 14 14
FT DISULFID 27 83
FT DISULFID 106 106
FT DISULFID 109 109
FT DISULFID 141 201
FT DISULFID 247 305
FT SEQUENCE 327 AA; 35940 MW; 3EDD811EF208E7A Cn64;
Query Match 29.6%; Score 1144; DB 1; Length 327;
Best Local Similarity 89.9%; Pred. No. 1.4e-72;
Matches 213; Conservative 7; Mismatches 9; Indels 8; Gaps 1;
QY 458 EPPYCPKPSCKIHTGDCCHAPPELLGGSVFLPPPKKIKLIMISRTPEVTCVVVDVSHED 527
DB 99 ESKYGP-----CSPAPEELGGSVFLPPPKPKDTLMISRTPEVTCVVVDVSHED 150
QY 528 PEVKFNMYVDGSEVHNAKTKPREQYNSTYRVVSVLTVLHDMWLNCKEYKCKVSNKALP 587
DB 151 PEVFNMYVDGSEVHNAKTKPREQYNSTYRVVSVLTVLHDMWLNCKEYKCKVSNKALP 210
QY 588 PIKTIKAKGQPPPEVYTLPPSPREMTKNQVSLTCLVKGFYPSDIAVWESNGQFENN 647

```



Db 211 SIETISKAKQPRPQVVTLPDSQEMTKNAVSLTLVKTFFVPSDIAYEVESNGQPPNN 276  
 QY 648 YKTIPTWIDSGFFLYSKLTVPKSKWQCNVPSCSVMHEALHNHYTKLSLSLSPCK 704  
 DB 271 YKTIPTWIDSGFFLYSKLTVPKSKWQCNVPSCSVMHEALHNHYTKLSLSLSPCK 327

RESULT 8  
 GC\_HABIT STANDARD; PRT; 323 AA.  
 AC P01870;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE Ig gamma chain C region  
 OS *Oryzotylagus cuniculus* (Rabbit).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryzologas  
 OX NCBI\_TaxID=9986;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 FX MEDLINE=8403940; PubMed=6319520.  
 RA Bernstein K F., Alexander C.R., Mage P.G.,  
 RT "Nucleotide sequence of a rabbit IgG heavy chain from the recombinant  
 RT F-1 haplotype".  
 RL Immunogenetics 18:387-397(1983).  
 RN [2]  
 RP SEQUENCE OF 1-128.  
 RX MEDLINE=76135469; PubMed=1243651;  
 RA Pratt D.M., Mole L.E.;  
 RT "Sequence studies on the constant region of the Fd sections of rabbit  
 RT immunoglobulin G of different allotype".  
 RL Biochem. J. 151:337-349(1975).  
 RN [3]  
 RP SEQUENCE OF 88-266 FROM N.A.  
 RX MEDLINE=83299917; PubMed=6194512.  
 RA Martens C.L., Moore K.W., Steinmetz M., Hood L., Knight K.L.;  
 RT "Heavy chain genes of rabbit IgG: isolation of a cDNA encoding gamma  
 RT heavy chain and identification of two genomic C gamma genes".  
 RL Proc. Natl. Acad. Sci. U.S.A. 79:6018-6022(1982).  
 RN [4]  
 RP SEQUENCE OF 132-161.  
 RX MEDLINE=70110015; PubMed=5461106;  
 RA Fuchter P.G., Jackson S.A., Mole L.E., Porter P.R.;  
 RT "Sequence studies of the Fd section of the heavy chain of rabbit  
 RT immunoglobulin G".  
 RL Biochem. J. 116:249-259(1970).  
 RN [5]  
 RP SEQUENCE OF 128-131 AND 155-122.  
 RA Hill P.L., Lebovitz H.E., Fellows R.E. Jr., Delaney R.;  
 RL (In) Killander J. (eds.);  
 RL Gamma globulins, Nobel symposium 3, pp.109-127, Almqvist and Wiksell,  
 RL Stockholm (1967).  
 CC -1- MISCELLANEOUS: REF.1 SEQUENCE HAS THE D12 ALLOTYPIC MARKER,  
 CC 104-THR, AND THE E14 MARKER, 185-THR. REF.3 HAS THE D11 AND E15  
 CC MARKERS AND REF.5 THE E15 MARKER.  
 CC  
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement. (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC  
 CC EMBL: M16426, AAA1289.1,  
 CC PIR: A02161, GHPR  
 CC HSSP: P01857, 1FCL.  
 CC InterPro: IPR003006; Ig\_MRC.  
 CC InterPro: IPR001597; Ig\_cl.  
 CC Pfam: PF00047; Ig\_3.  
 CC SMART: SM00407; Ig-L1\_2.  
 CC PROSITE: PS00240; Ig\_MHC; 1

FW Immunoglobulin chain C region.  
 FT NON-TER 1  
 FT VARIANT 104 104 T -> M (IN D11 MARKER).  
 FT VARIANT 185 185 A -> A (IN E15 MARKER).  
 FT CONFLICT 48 48 N -> E (IN REF. 2).  
 FT CONFLICT 71 71 V -> VPV (IN REF. 2).  
 FT CONFLICT 144 144 Q -> E (IN REF. 3 AND 4).  
 FT CONFLICT 173 173 N -> D (IN REF. 5).  
 FT CONFLICT 187 187 Q -> E (IN REF. 3 AND 5).  
 FT CONFLICT 201 201 N -> D (IN REF. 5).  
 FT CONFLICT 218 218 Q -> E (IN REF. 5).  
 FT CONFLICT 233 233 E -> Q (IN REF. 5).  
 FT CONFLICT 246 246 N -> D (IN REF. 5).  
 FT CONFLICT 256 256 E -> G (IN REF. 5).  
 FT CONFLICT 260 260 N -> D (IN REF. 5).  
 FT CONFLICT 266 266 N -> D (IN REF. 5).  
 FT CONFLICT 280 280 Y -> W (IN REF. 5).  
 FT CONFLICT 284 284 N -> S (IN REF. 5).  
 SQ SEQUENCE 323 AA; 35404 MW; 69E8AA118D579A8B CRC64;

Query Match 24.1%, Score 332, pH 2, Length 423;  
 Best Local Similarity 59.5%, Pct. 7.7e-58;  
 Matches 173; Conservative 40; Mismatches 48; Indels 30; Gaps 3;

QY 444 LKPLNA-----PNVDTGHNFAVINISSEPYFGEKPKCKDKTH----- 481  
 DB 33 LPEPVTVWNSGTLFNGVTFPSVQSSGLYSLSVSVSSQPVTCNVARPAINTKVD 92  
 QY 482 -----TC--PPCAPELLGGPSVLEPPKPKDILMISRTPEVTGVVVDVSHDEPKFN 533  
 DB 93 KIVASTSKKPLCPPPPELLGGPSVLEPPKPKDILMISRTPEVTGVVVDVSHDEPKFN 152  
 QY 534 WYVDGVFVNNAKTKRFEQYNIYKVVSVLTVLHDMVNGKRYCKVSKNKAIPAIKTI 593  
 DE 153 WYINQEVFIAPPLPEQCFNSTIPVNSTLPIHGWELPGREFKCKVINKALPAIIEKTI 612  
 QY 594 SKAKGPPPEPQVYVTLPPSKPEETKNQVSLTCLVKGFPESDIAYEVESNGQPPNNYKTTTP 653  
 DE 213 SKAPQCPPLPKVYTWGTPRELSVSVSTCTMNGFYSDISVWEKRGKGFADNYKTTTPA 272  
 QY 654 VLDSDGSFFLYSKLTVDKSRWQGNVPSCSVMHEALHNHYTKLSLSLSPCK 704  
 DB 273 VLDSDGSFFLYSKLTVDKSRWQGNVPSCSVMHEALHNHYTKLSLSLSPCK 323

RESULT 9  
 GC2\_CAVPO STANDARD; PRT; 329 AA.  
 AC P01862;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE Ig gamma-2 chain C region.  
 OS *Cavia porcellus* (Guinea pig).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Hystricognathi; Caviidae; Cavia.  
 OX NCBI\_TaxID=10141;  
 RN [1]  
 RP SEQUENCE OF 1-3.  
 RA Trischmann T.M.;  
 PL Submitted (Apr-1975) to the PIR data bank.  
 RN [2]  
 RP SEQUENCE OF 4-68.  
 RX MEDLINE=71058471; PubMed=5538606;  
 FA Birshwin B.F., Hussain Q.Z., Zebra J.J.;  
 ET "Structure of heavy chain from strain 13 guinea pig  
 RT immunoglobulin-g(2) 3 Amino acid sequence of the region around the  
 RT half-cysteine joining heavy and light chains.";  
 RL Biochemistry 10:18-25(1971).  
 RN [3]  
 RP SEQUENCE OF 69-133 AND 312-329.  
 RX MEDLINE=71058486; PubMed=5538616;



FT DOMAIN 783 1136 CYTOPLASMIC (POTENTIAL).  
 FT DOMAIN 43 106 IG-LIKE C2-TYPE DOMAIN 1.  
 FT DOMAIN 212 254 EGF-LIKE 1.  
 FT DOMAIN 256 301 EGF-LIKE 2.  
 FT DOMAIN 303 343 EGF-LIKE 3.  
 FT DOMAIN 370 424 IG-LIKE C2-TYPE DOMAIN 2.  
 FT DOMAIN 444 538 FIBRONECTIN TYPE-III 1.  
 FT DOMAIN 541 637 FIBRONECTIN TYPE-III 2.  
 FT DOMAIN 641 742 FIBRONECTIN TYPE-III 3.  
 FT DOMAIN 837 1116 PROTEIN KINASE.  
 FT NP\_BIND 843 851 ATP (BY SIMILARITY).  
 FT BINDING 868 868 ATP (BY SIMILARITY).  
 FT ACT\_SITE 977 977 HY SIMILARITY.  
 FT MOD\_RES 1005 1005 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).  
 FT CARBOHYD 84 84 N-LINKED (GLCNAC) (POTENTIAL).  
 FT CARBOHYD 159 159 N-LINKED (GLCNAC) (POTENTIAL).  
 FT CARBOHYD 501 501 N-LINKED (GLCNAC) (POTENTIAL).  
 FT CARBOHYD 594 594 N-LINKED (GLCNAC) (POTENTIAL).  
 FT CARBOHYD 707 707 N-LINKED (GLCNAC) (POTENTIAL).  
 SO SEQUENCE 1136 AA; 124953 MW; 8FF85804A041BB12 CRC64,  
 Query Match 22.8%, Score 882.5, Db 1, Length 1136,  
 Best Local Similarity 33.5%, Pred. No. 9.8e-54;  
 Matches 222; Conservative 89; Mismatches 250; Indels 101; Gaps 25;

Qy 16 LSGTVEGAMDLILNLSPLVSDAETSLTCT-----ASGWRPEHPETIGKDFEALMN 66  
 Db 15 LASHVGAADVLTLLADRLTEPQFFETLTVSGEAGAGSGDAMP--PLLEKDDRTVPT 72

Qy 67 QH--QDPLEVTQVTRWAKKVVWKP--FKASKINGAYFCHCRVGRGEAIRLTKMKMQQAS 123  
 Db 73 PRKQPP-----HIARGSSRVTVRFSQPSLLGVGSFVW-----GGGTRVLYVHNS-PRGAH 124

Qy 124 FLPATLITVDKGDNNVSEKVLIEEDAVIYKNGSFHISVPRHEVPD---ILEVHLPH 180  
 Db 125 LLPDKVTHVTKCDTAVLSARVRKKEQTDVIWKSNGSYFYTLDRHEADQGFILQ--LPN 182

Qy 181 ACPDGAIVYSAKYLIG-NLFTSATRLIVPRCEAKWKSPENHLCTATMNNVSVCHDTGEC 240  
 Db 183 VQPSGGIYSAIYLEASPLGSAFELLIVGCEAGWAGQCTKECPGCLHGGVCHDQGEQ 242

Qy 241 ICPHPCMCRTCEKACPLHTERTCKEPCSGEPYKSVYPLDPYSGSACATKWLQVNE 300  
 Db 243 VCPGFTSTWTELATPGLPFPQS--DELPATISAFPLIFPLPEYSGSASSWKSVALE 302

Qy 301 ACHPCFTGPGCKLDE--NN--HM--DRFG--GL--SP--WALQ--TEREH--IPMTPKIIVLDPHIE 360  
 Db 303 ACAPGRFCAVCHLQCCQDQNGTCTDRFSGGVCPSPSWHCHMCKESD---PILQLIDMVSELE 359

Qy 361 VNSCKENPI--CKASGWPPLTNEEMTLVKPGSTV--LHPKENHILHIESVALETHIRLPPD 418  
 Db 360 FNLDTMPINCAVGNFPFVSGSMELRKPSTVLLSTKAIVEPDP--TTAFEFVPRFALGD 418

Qy 419 SCWVWCNVNTVACVMPKPFNISKVLPKPLNAPNVDTHGNFAVINLSSRPPECEPKSCD 478  
 Db 419 SGLWECRVSTSGGDSPPFINVKVPVPVPLTAPLLAKQSPQLVVS-----464

Qy 479 KTHTCPPCPAPELLG--CP--SVTLFPKPKPDTLMSPTPTVTVVVSIE-----DPE 529  
 Db 465 -----PLVSTSGDGPFIASVRIH--YRPQDSTMWST-----IVVDPSENVTLMIIRPK 510

Qy 530 VKFNWYVGRVEVINAKTKPRFELVNS-----YRWVSVLIVLHJLWLN--KEY-----KCKVS 581  
 Db 511 T-----GYSVRVQLSRPEEGEGNAGWPPTLMTITPEPLLPKWLGLWIVE--PDPPLVVS 584

Qy 582 NKALPAP-----IEKTSKACQPPRPQVYTIHPSRFPMIKNCAVSPQLVKG--FYYS 632  
 Db 564 WSLPVPVPGPLVGGGFTLRLWCAFGDERRRNV--SSPLAFTAL-----LTSLPETGYVQL 616

Qy 633 DI 634  
 Db 617 DV 618

## RESULT 11

## TIE1\_HUMAN

ID TIE1\_HUMAN STANDARD; PRT: 1138 AA.  
 AC P35590;  
 DT 01-JUN-1994 (rel. 29, Created)  
 DT 01-JUN-1994 (rel. 29, Last sequence update)  
 DT 15-JUL-1999 (rel. 38, Last annotation update)  
 DE Tyrosine-protein kinase receptor TIE-1 precursor (EC 2.7.1.112).  
 GN TIE1 OR TIE.  
 OS Homo sapiens (Human).  
 OC Eukaryota, Metazoa, Chordata; Craniata, Vertebrata, Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID:9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA MEDLINE:92195316; PubMed:1312667;  
 RA Partanen J., Armstrong E., Mackelae T.P., Korhonen J., Sandberg M.,  
 RA Renkonen R., Knuutila S., Huebner K., Alitalo K.;  
 RT "A novel endothelial cell surface receptor tyrosine kinase with  
 RT extracellular epidermal growth factor homology domains.";  
 FL Mol. Cell Biol. 12:1638-1707(1992).  
 RN [2]  
 RP REVISIONS.  
 RA Partanen J.M.;  
 RL Submitted (JUL-1993) to the EMBL/GenBank/DBJ databases.  
 CC -!- FUNCTION: PROBABLE PROTEIN TYROSINE-KINASE TRANSMEMBRANE RECEPTOR.  
 CC -!- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein  
 CC tyrosine phosphate.  
 CC -!- SUBCELLULAR LOCATION: Type 1 membrane protein.  
 CC -!- TISSUE SPECIFICITY: SPECIFICALLY EXPRESSED IN DEVELOPING VASCULAR  
 CC ENDOTHELIAL CELLS.  
 CC -!- SIMILARITY: TO OTHER PROTEIN-TYROSINE KINASES IN THE CATALYTIC  
 CC DOMAIN.  
 CC -!- SIMILARITY: CONTAINS 2 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.  
 CC -!- SIMILARITY: CONTAINS 3 FIBRONECTIN TYPE III-LIKE DOMAINS.  
 CC -!- SIMILARITY: CONTAINS 3 EGF-LIKE DOMAINS.  
 CC This SWISS-Prot entry is copyrighted. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 EP EMBL: X60957; CAA43290.1;  
 EP HSSD: P11362; 1PGK.  
 DR MIM: 609222;  
 DR InterPro: IPR000561; EGF-like  
 DR InterPro: IPR000719; Euk\_kinase.  
 DR InterPro: IPR000461; FN\_III.  
 DR InterPro: IPR003006; Ig\_MHC.  
 DR InterPro: IPR001245; Tyr\_kinase.  
 DR Pfam: PF00008; EGF\_2.  
 DR Pfam: PF00041; In3; 3.  
 DR Pfam: PF00047; Ig\_2.  
 DR Pfam: PF00664; fkinase; 1.  
 DR PRINTS: PR00109; TYRKINASE.  
 DR SMART: SM00181; PGP; 2.  
 DR SMART: SM00060; FN3; 2.  
 DR SMART: SM00219; TyroK; 1.  
 DR PROSITE: PS00107; PROTEIN\_KINASE\_ATP; 1.  
 DR PROSITE: PS00109; PROTEIN\_KINASE\_TYP; 1.  
 DR PROSITE: PS00011; PROTEIN\_KINASE\_DOM; 1.  
 DR PROSITE: PS00022; EGF\_1; 3.  
 DR PROSITE: PS01186; EGF\_2; 3.  
 KW Receptor; Tyrosine-protein kinase; Transferase; Signal; ATP-binding;  
 KW Repeat; EGF-like domain; Transmembrane; Immunoglobulin domain;  
 KW Glycoprotein; Phosphorylation; Multigene family.  
 FT SIGNAL 1 24 POTENTIAL.  
 FT CHAIN 25 1138 TYROSINE-PROTEIN KINASE RECEPTOR TIE-1.

```
FT DOMAIN 25 759 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 760 784 POTENTIAL.
FT CYTOPLASMIC (POTENTIAL).
FT DOMAIN 785 1138
FT DOMAIN 43 105 IG-LIKE C2-TYPE DOMAIN 1.
FT DOMAIN 214 256 BCF-LIKE 1.
FT DOMAIN 258 363 BCF-LIKE 2.
FT DOMAIN 365 445 BCF-LIKE 3.
FT DOMAIN 372 426 IG-LIKE C2-TYPE DOMAIN 2.
FT DOMAIN 446 540 FIBRINOLYTIC TYPE-III 1.
FT DOMAIN 543 619 FIBRINOLYTIC TYPE-III 2.
FT DOMAIN 643 744 FIBRINOLYTIC TYPE-III 3.
FT DOMAIN 849 1118 PROTEIN KINASE.
FT NP_BIND 845 853 ATP (BY SIMILARITY).
FT BINDING 876 870 ATP (BY SIMILARITY).
FT ACT_SITE 979 979 HY SIMILARITY.
FT CARBOHYD 83 83 N-LINKED GLYCAC. (POTENTIAL).
FT CARBOHYD 161 161 N-LINKED GLYCAC. (POTENTIAL).
FT CARBOHYD 503 503 N-LINKED GLYCAC. (POTENTIAL).
FT CARBOHYD 546 546 N-LINKED GLYCAC. (POTENTIAL).
FT CARBOHYD 709 709 N-LINKED GLYCAC. (POTENTIAL).
FT MOD_RES 1067 1067 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
SQ SEQUENCE 1138 AA: 125089 MW: 38428633678058A1 CRC64.

Query Match 22.6%, Score 874.5, DB 1, Length 1138,
Best Local Similarity 31.2%, Pred. No. 3.5e-53,
Matches 214, Conservative 99, Mismatches 23, Gaps 19.

QY 14 LLSLSIVGAMDLILINSIPVSDAFTSLICL-----ASQRRPHEPITIGRDEAL 64
DB 13 LFLASHVGAADLTLLANRLTDPREFELVCGENANRASPAPPP-PLLEKQKIV 70

QY 65 MNDHQIPLVTOVQVTRWAKKVVWKR-IRASKINGAYPCGVVGVGEALIRPTMKRQCA 123
DB 71 RTRPQHPLEP-----LAPNGSHQVTHPEKSPKSLVGVSEVCGACAPRIRVIVHNSGCAH 126

QY 124 FLPATLTMTVDKQDNVTSFKVLTKEDAVLYKNGSHVSPRIEVPD---LLEVLHP 180
DB 127 LLDKVTIIVNKGDIYAVISARVHKIKQTVLWKSNGSYFLDWEHAGQKFLQ--LPI 184

QY 181 AQDQAGVVSARYTQNTPTSAFTPIVPPCFACKWSPQNNHICVACMNNVGVDFPDQ 240
DB 185 VUPSSSISALYLEASPLGSAFFELVKGAGVAPWSP-LKEPGLVGV-ILHGLSE 244

QY 241 ICQPCFMGRICQKACGLHIFQICNFKSCQVQCKSVVFCILPDPYQSCALGWKGLQ 300
DB 245 VCPPTGTTRCQACRQGFQSCQCPQCGLSQCKGLIFCLPDPYQSCQSGWKGSCQ 304

QY 301 ATRPFPYAPCKLRKSPNNKTEMTREGLQGLSPWGLWLEPEIIPPMTPKIVILPQ 360
DB 305 ACAQCHFGADCHLOFCQNGGICDNFSCQVCPQSGWGHVCHCKSD--RIPOLLN 361

QY 361 VNSCKNFI-CKASQWPIPIPPMTLVKPDQTVHPKISNHTDFSVAILPTIRHILP 419
DB 362 FLELTMPFINCAANPFPVRSIGIELRPGDQTVLLSTKAIVEPKTTAEFEVPLVL 421

QY 420 GWWVSVNIVACMVKPKNISVKVLPKLNALNVLDIHNHFAVIN ISSPPHYGE-PKS 476
DB 422 GWKCHVPSISQCGDSRPAKAVVGVVVAALATLTKLSPGVVSPIVS---PSL 478

QY 477 QDKTHITPCPCAPPELLGGSPSELEPPKPKDITMSPTPEVTVVVVVVSH----- 525
DB 479 IVRLHYRIPQDSIMDW-----SLIVDPSENVTLMNLRPKTGYSVRVQLSRPG 534

QY 526 -----EPEVAPF-----NWKVQCVF-----VHNKATKPRE 550
DB 535 PTLMTDCKPEQLQWLEGVLEWGTDRUKVSWSLPLVGPPLVGPGLPLMDGPPGEP 554

QY 551 EGYNSTIYRVSVLTVL-----HQDQWNGKRYKCKVSNKALPAPTEKTLISKAG 606
DB 595 ENVSSIPQANTALLIGLPGHYQ LDDQVHCHILIGPASD-----PAHV 637

QY 607 TLFPSS-- REEMTKNVSILT 623
```

```
DB 638 LLPPSGPPAPRHILHAGALSDSEIQLT 663

RESULT 12
GC3_MOUSE STANDARD: PRT: 329 AA:
ID GC3_MOUSE
AC P22436;
DT 01-AUG-1991 (Rel. 19, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Ig gamma-3 chain C region, secreted form.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
SEQUENCE FROM N.A.
MEDLINE=85027161; PubMed=8692053;
RA Wells T.A., Word C.T., Pimm D., Der-Balan G.P., Martinez H.M.,
RA Tucker P.W., Blattner P.R.;
RT "Structural analysis of the murine IgG3 constant region gene";
RL FMBR J. 3:2041-2046(1984).
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed, usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR FMBL; J00451; ; NOT_ANNOTATED_CDS.
DR PIR; R02156; G3MSC.
DR HSSP; P01857; 1FC1.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003597; Ig_C1.
DR InterPro; IPR003600; Ig_Like.
DR Pfam; PF00047; Ig; 3.
DR SMART; SM00410; Ig_Like; 1.
DR SMART; SM00407; ICG1; 2.
DR PROSITE; PS00290; IG_MHC; 1.
KW Immunoglobulin domain; Immunoglobulin C region; glycoprotein;
KW Transmembrane; Alternative splicing.
FT NON_TER 1
FT DOMAIN 1 97 CH1.
FT DOMAIN 98 113 HINGE.
FT DOMAIN 114 223 CH2.
FT DOMAIN 224 327 CH3.
SQ SEQUENCE 329 AA: 36228 MW: 45827174182BAD6 CRC64;

Query Match 22.2%, Score 857.5, DB 1, Length 329,
Best Local Similarity 53.5%, Pred. No. 1.2e-52,
Matches 161, Conservative 41, Mismatches 62, Indels 37, Gaps 3;

QY 409 FTHRIILPDSCW-----VGSNIVACMVKPKNISVKVLPKLNAPNVLDIHNHFAVI 463
DB 61 YSLSELVVPSSVWSTWISGVICNVAFHASKTFLIRIPRI--PKPSTPPG----- 108

QY 464 NISSEPFSEPKSDEKTHITPCPCAPPELLGGSPSELEPPKPKDITMSPTPEVTVVV 524
DB 109 -----SSCPDGNILGGVSFVLPKPKDIALISLTKVTCVVVDV 148

QY 524 SHEDPEVFNHYVVDGVVHNNAKTKIPRQYNSITRVVSVLTVLHQDWLNGKEYKCKVSNK 583
DB 149 SEDRDEIVSVFVFNKFIITAMTQPEAQYNSITPEVVSALPFIQDQWMPKETFCVKNK 208

QY 584 ALPAPTEKTLISKAGQPREPVYTIIPPSKREMTKNQVSLICLVKGFYPSDIADVHNSGO 643
DB 209 ALPAPTEKTLISKAGQAPGVYTIIPPSKREMTKNQVSLICLVNIPFEALSVHNSG 268

QY 644 PENNYKITPPVLLSDGSEFLYSKLTIVKQSEWQGVNVSQSYMHEALHNIVTKQSLSPG 703
```

```

Db 269 LEQYKNTPTILSDGTYFYYSKLIVTDWISLQCEIFTCVSWHEALHNHHTKNSRSPG 328
QY 704 K 704
Db 329 K 329

RESULT 14
GCB_RAT
ID GCB_RAT STANDARD: PRT: 333 AA.
AC P20761:
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 15 JUL 1999 (Rel. 38, Last annotation update)
DE Iq gamma-2B chain C region.
OS Rattus norvegicus (Rat).
OC Eukaryota, Metazoa, Chordata, Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia, Sciurognathi; Muridae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89232738; PubMed=3140946;
RA Bruegemann M.;
RL "Evolution of the rat immunoglobulin gamma heavy-chain gene family.";
DR Gene 74:473-482(1988).
DR PIR: PS0018; PS0018.
DR HSSP: P01842; 7FAB.
DR InterPro: IPR003006; Iq_MHC.
DR InterPro: IPR003597; Iq_CL.
DR Pfam: PF00047; Iq; 3.
DR SMART: SM00410; IG_Like; 1.
DR SMART: SM00407; IGCL; 2.
DR PROSITE: PS00230; IG_MHC; 1.
KW Immunoglobulin domain; Immunoglobulin C region.
FT NON_TER 1 1
FT DISULFID 15 15 INTERCHAIN (WITH A LIGHT CHAIN).
FT DISULFID 27 80
FT DISULFID 106 106 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 109 109 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 112 112 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 115 115 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 147 207
FT DISULFID 253 311
SQ SEQUENCE 333 AA: 30499 MW: 55F86e4d4b460Ab (R64;

Query Match
Best Local Similarity 43.4%; Pred. No. 3.2e-52.
Matches 163, Conservative 47, Mismatches 73, Indels 47, Gaps 3;

QY 415 LPDSGWVCSVNTVAGWKEKPNISVKVLPKPLNA-----PNVIDTGHNF 460
Db 11 LAOCGGTSTSTVTLGLVAGYF-----PEPVTVTWNSGALSSDVHTFPAVLQSLYR 63
QY 461 AVINISSEPEVGEKSCDKTH-----TCPPCPAPELIGG 494
Db 64 LTSVSTSTSPSQTNTVAHPASSTKVDKKEPNNGIGHKPTPTPTCHKCPVELIIG 123
QY 495 PSVFLFPKPKDITIMISRTPEVTCVVDVDSHEDPEVKFNWYDGVVHNAKTKPREFOYN 554
Db 124 PSVFIFFPKPKDILLISQNAKVICVVDVSEEPDQVFSWVNVNVEHTAUPPEBYN 184
QY 555 STYRVSVVLVLQDWLNGEYKCKVSNKALPAPIEKTISKAKGQPREPOVYTLPPSRRE 614
Db 184 STFRVVSALPIQHDWMSGKEFKCKVNNKALPSPIETKIIISKPKLVKKIQQVVMGDPTEQ 243
QY 615 MTKNQVSLTCLVKGFYPSDIAVEESGQPENNYKTTTPVLDSDGSFFLYSKLTVDKSRW 674
Db 244 LTEGVTSLTCLTSGLFENDIGVETWSNGHIEKKNKNTPEVWDSGSEFFMYSKLNVSRW 303
QY 675 QQGNVSCSVNHEALHNHHTKNSRSPG 704

```

```

Db 304 DSRAPVCSVWHEGLRNHHVKSISNPPCK 333

RESULT 14
GCB_MOUSE
ID GCB_MOUSE STANDARD: PRT: 398 AA.
AC P03987:
DT 23-OCT-1984 (Rel. 02, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 15 JUL 1999 (Rel. 38, Last annotation update)
DE Iq gamma-3 chain C region, membrane-bound form.
OS Mus musculus (Mouse).
OC Eukaryota, Metazoa, Chordata, Vertebrata; Euteleostomi;
OC Mammalia; Eutheria, Rodentia, Sciurognathi, Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85027161; PubMed=6392053;
RA Wells J.A., Ward C.J., Pimm D., Der-Balan G.P., Martinez H.M.,
RA Tucker P.W., Blattner F.R.;
PT "Structural analysis of the murine IgG3 constant region gene.";
PL EMBO J. 3:2041-2046(1984).
RN [2]
RP SEQUENCE ON 328-398 FROM N.A.
RX MEDLINE=84041483; PubMed=6314258;
RA Komaromy M., Clayton L., Rogers T., Robertson S., Kettman J.,
RA Wall R.;
RT "The structure of the mouse immunoglobulin in gamma 3 membrane gene
segment.";
RL Nucleic Acids Res. 11:6775-6785(1983).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
EMBL: J00451; AAB59655.1; .
EMBL: V01526; CAA24767.1; ALT_SFQ.
PIR: A02155; G3MSM.
HSSP: P01857; 1FCL.
InterPro: IPR003006; Iq_MHC.
InterPro: IPR003597; Iq_CL.
InterPro: IPR003600; Iq_Like.
Pfam: PF00047; Iq; 3.
SMART: SM00410; IG_Like; 1.
SMART: SM00407; IGCL; 2.
PROSITE: PS00230; IG_MHC; 1.
KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein;
KW Transmembrane; Alternative splicing.
FT NON_TER 1 1
FT DOMAIN 1 97 CH1.
FT DOMAIN 98 113 HINGE.
FT DOMAIN 114 233 CH2.
FT DOMAIN 234 327 CH3.
FT TRANSMEM 346 362 POTENTIAL.
FT DOMAIN 363 398 CYTOPLASMIC (POTENTIAL).
FT CONFLICT 333 342 E -> G (IN REF. 2).
FT CONFLICT 342 343 E -> Q (IN REF. 2).
FT CONFLICT 388 388 P -> F (IN REF. 2).
SQ SEQUENCE 398 AA: 43929 MW: C7F264B50A41B95 CRK64;

Query Match
Best Local Similarity 21.9%; Score 84e-5; DB 1; Length 398;
Matches 159; Conservative 41; Mismatches 62; Indels 37; Gaps 3;

QY 409 FTHITLPPDSGVW-----VCSVNTVAGWKEKPNISVKVLPKPLNAPWINTGHNAVI 463
Db 61 YLSLSIVTPSSSTWPSQTVCNVAHPASSTKVDKKEPNNGIGHKPTPTPTCHKCPVELIIG 108

```



Genre version 4 5  
Copyright (c) 1993 - 2000 Component Ltd.

OM protein - protein search, using sw model

Run on: September 23, 2002, 22:09:44 Search time 72.16 Seconds  
(without alignments)  
1687.755 Million cell updates/sec

Title: US-09-733-764-2  
 Perfect score: 386A  
 Sequence: 1 MDSLASLVLCVSLILGGTV MIEALINHYTKSISLSPGK 704

Scoring table: `BLOGSUM2`  
`GapOp 10.0` , `GapExt 0.5`

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

```
Database :
SPTEMP_L9_*
1: sp.archaea.*
2: sp.bacteria
3: sp.fungi.*
4: sp.human.*
5: sp.invertebr.*
6: sp.mammal.*
7: sp.mhc.*
8: sp.oranell.*
9: sp.phage.*
10: sp.plant.*
11: sp.protist.*
12: sp.virus.*
13: sp.vertebra.*
14: sp.unclass.*
15: sp.virus.*
16: sp.bacteria
17: sp.archaea
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARY

Result No.	Query %			Length	DB	ID	Description
	Match	Score					
1	2293	5	59.3	1082	11	Q9QM24	Q9QM24 rattus sp.
2	1303	4	33.7	701	4	Q9QF38	Q9QF38 homo sapien
3	916	5	23.7	337	6	Q95M34	Q95M34 equus caball
4	9015	23.3	1116	13	Q73731	Q73731 bruchydanic	
5	849	21.9	463	1	Q991C4	Q991C4 mus muscul	
6	845	21.8	437	1	Q9R1A4	Q9R1A4 mus muscul	
7	825	21.3	473	11	Q9D814	Q9D814 mus muscul	
8	817	21.1	473	11	Q99125	Q99125 mus muscul	
9	816	21.1	468	11	Q99L31	Q99L31 mus muscul	
10	788	20	4	473	11	Q91705	Q91705 mus muscul
11	360	9.3	597	4	Q96B89	Q96B89 homo sapien	
12	359	9.3	375	4	Q9HS71	Q9HS71 homo sapien	
13	359	9.3	597	4	Q9BQ88	Q9BQ88 homo sapien	
14	359	9.3	597	4	Q9BU10	Q9BU10 homo sapien	
15	355	9	613	4	Q96EY8	Q96EY8 homo sapien	
16	349	9	614	4	Q96GAK	Q96GAK homo sapien	

## ALIGNMENTS

```

RESULT 1
SEQUENCE FROM N.A.
MEDLINE-3275658; PubMed-7684830;
Maisonpierre P.C., Goldfarb M., Yancopoulos G.D., Gao G.J
"Distinct rat genes with related profiles of expression define a 116
receptor tyrosine kinase family.";
Oncogene 8:1631-1637(1994).
HSSP; P11362; IFGK.
InterPro: IPR000561; EGF-like.
InterPro: IPR000719; Euk_pkinase.
InterPro: IPR003961; FN_III
InterPro: IPR002649; laminin_EGF.
InterPro: IPR001245; Tyr_pkinase.
Pfam: PF00041; fn3_3.
Fam, PF00069; pkinase; 1.
PRINTS: PP00011; EGF_LAMININ.
PRINTS: PK00109; TYRKINASH.
SMART: SM00180; EGF_Lam; 1.
SMART: SM00003; EGF_Like; 1.
SMART: SM00060; FN3_3.
SMART: SM00219; TYRK3_1.
PSSITE: PSC0023; EGF_1; UNKNOWN_3.
PROSITE: PS01165; EGF_2; 3.
PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
Prosite: PS00109; PROTEIN_KINASE_TYP; 1.
ATP-binding; EGF like domain; Glyceraldehyde-3-phosphate

```







```

Db 109 R-NSTGRKLYIKYMLQEAALFPESLITIVNOGGINISYSRLYSPEDTVIHKNCHPEHS 167
QY 165 VPHFEVPTLLEVLPLNAPQIP-AWYSAPYTGQNIPTSAFTPLIVPPEACQKWPPEPNHL 223
Db 168 SPEKISDTHVPTNKAENHRIYAIYISAAPSSAAITPLIVRSCFAGFWPNFTES 227
QY 224 CTAAMNNVTHVPTNKAENHRIYAIYISAAPSSAAITPLIVRSCFAGFWPNFTES 282
Db 228 CFCANGWQETTELAVVTPHGHVAVVGESEFCAQKEP VIGWVALVPLP 285
QY 283 DEYGSICATIGWGLNENRTHNFEVPLKLPSTNNEMDPPEGLNAPWVHGLP 342
Db 286 DEYGSICATIGWGLNENRTHNFEVPLKLPSTNNEMDPPEGLNAPWVHGLP 344
QY 343 EGIPMTKPIVLPDIHVEVNSG-KENPTCKASGMPLEPTNEEMTLVKPDGVLLHPKDFNHT 401
Db 344 AD---SSPVLSILRD-VEINTGVLSVNSGASGPPALHQPDTLITANPTTTAAVTHTL 399
QY 402 DHFSVAITFTTHLPPDSGVVCSVNTVAGWVEKPFNISVKVLKPLNAPNVIDTGHNFA 461
Db 406 NGQSVFVAVGVVRS-SAPGWPVNNNTIIMAVLEFTEVVKVPPPPNPPVLQSSPPHL 454
QY 462 VINISSEPYEGE 473
Db 460 LLLINTEPYSGD 471

RESULT 5
Q99LC4
ID Q99LC4 PRELIMINARY; PRT: 463 AA.
AC Q99LC4;
DT 01-JUN-2001 (Tremblrel, 17, Created)
DI 01-JUN-2001 (Tremblrel, 17, Last sequence update)
DI 01-DEC-2001 (Tremblrel, 19, Last annotation update)
DE SIMILAR TO P1KEN CNA 141-066-000 GENE
OS Mus musculus (Mouse)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RC Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases
DR EMBL: AC003439; AA003435.1;
DR ISSP: P01842; 7FAB.
DR InterPro: IPR003599; Iq.
DR InterPro: IPR003600; Iq_c1.
DR InterPro: IPR003006; Iq_MHC.
DR Pfam: PF00047; Iq_V.
DR SMART: SM00409; Iq; 3.
DR SMART: SM00409; Iq; 2.
DR SMART: SM00406; Iq; 1.
DR SMART: SM00410; Iq_Like; 1
DR PROSITE: PS00290; Iq_MHC; UNKNOWN_1.
SQ SEQUENCE 463 AA: 51007 MW: 54874.056K30783 CRC64:

Query Match 21.4%; Score 849; DH 11; Length 463;
Best Local Similarity 45.4; Pred No. 1.2e-69;
Matches 160; Conservative 55; Mismatches 65; Indels 44; Gaps 6;

```

```

QY 415 LPPDSGVVCSVNTVAGWVEKPFNISVKVLKPL-----NAPNVDTGHNF----- 460
Db 150 LAPGSAAGTNSMVTIGCLVKGYF-----PEPVITVNSGSLSSGVVHTPAVLQSDLYT 202
QY 461 --AVINISSEPYEFPKSKTHF-----CPP--CPAPELLGQPSVFLP 500
Db 203 LSSSVLVPSSTWSPSEIVTCNVAHFASSTKVDKIVPDDGCKPCICITVPEV---SSVFIF 259
QY 501 PPKPKDLTILTPKVTICVVVWVSHEDPEKFNWYVAGVGVHNAKINPREQVNSIYRVV 560

```

```

Db 260 PPKPKDVLITLTPKVTICVVVWVSHEDPEKFNWYVAGVGVHNAKINPREQVNSIYRVV 319
QY 561 SVLTVLHGLWLNKREYKCKVSNKALPAPLEKTSKAKQEPPEFQVYTLPPSPREMTKNQV 620
Db 320 SEIPMHQDLWLNKREKCKVSNKALPAPLEKTSKAKQEPPEFQVYTLPPSPREMTKNQV 379
QY 621 SVLTVLHGLWLNKREYKCKVSNKALPAPLEKTSKAKQEPPEFQVYTLPPSPREMTKNQV 680
Db 380 SLTCLVKKFEYFSDTAVENSNQFENNYKTTTPVLDSDPSEFELYSKLIVKSWKSGNVF 680
QY 681 SCVSMHEALHNHYTKSLSLSLSPGK 704
Db 440 TQSVLHGLWLNKREKCKVSNKALPAPLEKTSKAKQEPPEFQVYTLPPSPREMTKNQV 449

RESULT 6
Q99IA4
ID Q99IA4 PRELIMINARY; PRT: 437 AA.
AC Q99IA4;
DT 01-MAY-2000 (Tremblrel, 13, Created)
DI 01-MAY-2000 (Tremblrel, 13, Last sequence update)
DI 01-DEC-2001 (Tremblrel, 19, Last annotation update)
DE GAMMAL HEAVY CHAIN OF MAB7 (FRAGMENT)
OS Mus musculus (Mouse)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Wilco K.G.; Yu X.; Kramoddoullah A.K.M.; Misra S.;
RC Cloning of cDNAs encoding for anti-white pine blister rust monoclonal
RT antibody (Mab 7, its light and heavy chains) and construction of a
RT single chain antibody (scFv).
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF152372; AA040343.1;
DR ISSP: P01842; 7FAB.
DR InterPro: IPR003600; Iq_Like.
DR InterPro: IPR003006; Iq_MHC.
DR InterPro: IPR003599; Iq_c1.
DR Pfam: PF00047; Iq; 4.
DR SMART: SM00406; Iq; 1.
DR SMART: SM00410; Iq_Like; 2.
DR PROSITE: PS00290; Iq_MHC; UNKNOWN_1.
FT NON_TER 1
FT NON_TER 437
SQ SEQUENCE 437 AA: 48142 MW: 503A7B3EE7D647C CRC64:

```

```

Query Match 21.8%; Score 845; DH 11; Length 437;
Best Local Similarity 49.1%; Pred No. 2.5e-69;
Matches 159; Conservative 56; Mismatches 65; Indels 44; Gaps 6;
QY 415 LPPDSGVVCSVNTVAGWVEKPFNISVKVLKPL-----NAPNVDTGHNF----- 460
Db 134 LAPGSAAGTNSMVTIGCLVKGYF-----PEPVITVNSGSLSSGVVHTPAVLQSDLYT 176
QY 461 --AVINISSEPYEFPKSKTHF-----CPP--CPAPELLGQPSVFLP 500
Db 177 LSSSVLVPSSTWSPSEIVTCNVAHFASSTKVDKIVPDDGCKPCICITVPEV---SSVFIF 243
QY 501 PPKPKDLTILTPKVTICVVVWVSHEDPEKFNWYVAGVGVHNAKINPREQVNSIYRVV 560
Db 234 PPKPKDVLITLTPKVTICVVVWVSHEDPEKFNWYVAGVGVHNAKINPREQVNSIYRVV 294
QY 561 SVLTVLHGLWLNKREYKCKVSNKALPAPLEKTSKAKQEPPEFQVYTLPPSPREMTKNQV 620
Db 294 SLTCLVKKFEYFSDTAVENSNQFENNYKTTTPVLDSDPSEFELYSKLIVKSWKSGNVF 354
QY 621 SVLTCLVKKFEYFSDTAVENSNQFENNYKTTTPVLDSDPSEFELYSKLIVKSWKSGNVF 680
Db 384 SLTCLVKKFEYFSDTAVENSNQFENNYKTTTPVLDSDPSEFELYSKLIVKSWKSGNVF 414
QY 681 SCVSMHEALHNHYTKSLSLSLSPGK 704

```







```

DE 426 FSAWGEASL:EDWNSGERETGIVHIDLPSPKQJLSEPKVALHPPCVLLPPAPAEQL 486
QY 616 T-KNWSLSLTLKGVSPGDIHAWHWSNCP--ENNYKTIIPVLD--SNGSPFLYSKLTVLD 670
DB 486 NUKESATITCLTVGSPALVFWVMWCRGQPLSPKQVTSAPMPQAPGRYFAHSILTVS 545
QY 671 KSRWQGNVFSQVSMHEALHNIYOKSLSPGK 704
DB 546 EEWNTGTYTCVVAHFAALPNHVTIRKIDKSIQK 579

RESULT 14
Q96B010 PRELIMINARY: PRT: 597 AA.
AC Q96B010
DT 01-JUN-2001 (TREMURel 17, Created)
DT 01-JUN-2001 (TREMURel 17, Last sequence update)
DT 01-DEC-2001 (TREMURel 19, Last annotation update)
DE HYPOTHETICAL 65.3 KDA PROTEIN.
OS Homo sapiens (Human).
OC Eukaryota, Metazoa, Chordata, Vertebrata, Euteleostomi,
OC Mammalia, Theria, Primates, Catarrhini, Hominoidea, Homo.
OX NCRI_TaxID=9606;
RN [1]
SEQUENCE FROM N.A.
RA Strausberg R.;
RI Submitted (FEB-2001) to the EMBL/GenBank/DDBJ databases.
DR EMBL: AC002963; AA002963.1;
DR BSSP: P01825; 7FAB.
DR InterPro: IPR003599; Iq_1.
DR InterPro: IPR003600; Iq_1.1ko.
DR InterPro: IPR003606; Iq_1.1ko.
DR InterPro: IPR003596; Iq_1.
DR Pfam: PF00047; Iq; 5.
DR SMART: SM00403; Iq; 2.
DR SMART: SM00407; Iq; 1.
DR SMART: SM00410; Iq; 1.
DR PROSITE: PS00290; Iq_MHC; UNKNOWN_3.
KW Hypothetical protein.
SQ SEQUENCE 597 AA: 45.74 MW: 214647.07 E=5.8e-10664.

```

Query Match 9.3%; Score 359; DH 4; Length 597;  
 Best local similarity 34.1%; Pred. No. 3e-24;  
 Matches 73; Conservative 48; Mismatches 85; Indels 8; Gaps 6;

```

QY 497 VLEPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNKYVDGVVEHNAKTKPREEQYNST 556
DB 368 VFAIDPVS-FASITLTKSTIKITCLVLDITYD-SVTISWTRONGRAVKTHTNISSEHPNAT 425
QY 557 YRVVSVLTVLIQDGLNGKEYKIKVSNKALPAPIERTISAKKQGP-REPQVYTLPPSPREEM 615
DB 426 FSAWGEASL:EDWNSGERETGIVHIDLPSPKQJLSEPKVALHPPCVLLPPAPAEQL 485
QY 616 T-KNWSLSLTLKGVSPGDIHAWHWSNCP--ENNYKTIIPVLD--SNGSPFLYSKLTVLD 670
DB 486 NUKESATITCLTVGSPALVFWVMWCRGQPLSPKQVTSAPMPQAPGRYFAHSILTVS 545
QY 671 KSRWQGNVFSQVSMHEALHNIYOKSLSPGK 704
DB 546 EEWNTGTYTCVVAHFAALPNHVTIRKIDKSIQK 579

```

```

RESULT 15
Q96EY0 PRELIMINARY: PRT: 613 AA.
AC Q96EY0
DT 01-DEC-2001 (TREMURel 19, Created)
DT 01-DEC-2001 (TREMURel 19, Last sequence update)
DT 01-DEC-2001 (TREMURel 19, Last annotation update)

```

```

DE UNKNOWN {PROTEIN FOR M30:20337}.
OS Homo sapiens (Human).
OC Eukaryota, Metazoa, Chordata, Vertebrata, Euteleostomi,
OC Mammalia, Theria, Primates, Catarrhini, Hominoidea, Homo.
OX NCRI_TaxID=9606;
RN [1]
SEQUENCE FROM N.A.
RA Strausberg R.;
RI Submitted (JUL-2001) to the EMBL/GenBank/DDBJ databases.
DR EMBL: AC011857; AA011857.1;
SQ SEQUENCE 613 AA: 67.73 MW: 31214.04 E=64.21E7 CRC64;

```

Query Match 9.2%; Score 355; DH 4; Length 614;  
 Best local similarity 21.6%; Pred. No. 7.2e-24;  
 Matches 56; Conservative 92; Mismatches 249; Indels 246; Gaps 48;

```

QY 34 LVSDAET-SLTGASGWPPEPITTPRFFALMNOHPPLEVTQVTRMAKKVWKEK 92
DB 30 LVKPSHLISLIVSS-----QSISSV-----WS--WIRQP 59
QY 93 ASKINGATFCHGCPVPCPAIPPTMPMPQDASHLPATLTMVVKCNVNISSPKVLIKEED 152
DB 60 AGK--GLEWIG-----RITSGTNYNPSLSKSRVMSVDTSKN-QFSLKLSSTAAAD 108
QY 153 AVLYKNGSPFHSVORHEVDHLEVHPHQAQDQACVVSARVIG-CNIFISAFRLVVRK 211
DB 109 TAVYYCAS-----QWELPVLGLFYWGUGLIVVS-----Y 148
QY 212 FAGKWPCEMILCTACMNGVHEETGECIDPFPMPTPEKA'ELITETGT'VEKESNQ 271
DB 139 -----SOSASAPILFPLVSCF-----NNSDSEVVA 164
QY 272 FCKSVVFCIPHYGSCATQWGLQCNF-ACHQPF-----Y 407
DB 165 VGLAQDF-LPD-----STF-NKYNNSSGSS-3GPPSVI6GKYAATSQVLLSKVMQ 219
QY 308 GPI-----CKLPGCONNGEMCHPFGCCGSPGCGIQCEREGIPRMPTKIVLIDPHIEV 361
DB 220 GTDHEHWCKVQHNGNKE-----KNVP--LIVIAFLPHKVSFV 256
QY 362 NSGKFNDCYASG WFLTNEEMTLVKPDGTVLIPKDFNHTDHF 404
DB 257 PRDGHFGNPKRSKLIICVATGFSRQIQVSWIRGKQVGSVITDQVQALAKESGPTTYK 416
QY 405 SVAIFTHRLPPISGVVCSVNTVAGNVEKPIETSVKVLKPLNAPNVIIDIGHNFAVIN 464
DB 317 VSTILTIKESDWSLQSMFTCRVDH-----RGLTFQONASSMCVPDQDTAL-- 461
QY 465 ISSEYPFGPKSCHTHICPPCPAPELIGGPSVFLPHPKPDITLMSKTPETVTCVVVWS 524
DB 362 -----KVFALDPS-FASITLTKSTIKITCLVLDIT 489
QY 525 HEDPEKFNWYVDGVVEHNAKTKPREEQYNSTYRVVSVLTVLIQDGLNGKEYKIKVSNKA 584
DB 390 TYD-SVTISWTRONGRAVKTHTNISSEHPNATFSAVGHASICHEDWNSGRPTCTVTHTD 448
QY 585 LPAPITKTIKAKGQP-REPQVYTLPPSPREEM--KNWSLSLTLKGVSPDIHAWHWSN 642
DB 449 LPSPLKQTIISRPKGVALHPPDVYLLPPAREQLNLRSATITCLVTFGSPALVFWVMW 508
QY 643 QP--ENNYKTIIPVLD--SNGSPFLYSKLTVKSKWQGNVFSQVSMHEALHNIYOKSL 698
DB 509 QPISPEKVTISAPMPQAPGRYFAHSILTVSEFHWNTGTYTCVVAHFAALPNHVTIRK 548
QY 699 SLS 701
DB 569 DKS 571

```

Search completed: September 23, 2002, 22:12:59

Job time: 195 sec

---





GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

QM protein - protein search, using sw model

Pub. on: September 23, 2002, 22:11:29, Search time: 00:04 seconds  
(without alignments)  
1699 962 Million cell updates/sec

Title: US-09-733-764-2\_copy\_1\_472  
Perfect score: 2605  
Sequence: 1 MDSASLVLCVSLISGV VITSHNPNVINISSEYFG 472

Scoring table:  
Gapop 10.0, Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database: 1: /SIDSL/gcdata/geneseq/geneseq-emb1/AA1980.DAT:  
2: /SIDSL/gcdata/geneseq/geneseq-emb1/AA1981.DAT:  
3: /SIDSL/gcdata/geneseq/geneseq-emb1/AA1982.DAT:  
4: /SIDSL/gcdata/geneseq/geneseq-emb1/AA1983.DAT:  
5: /SIDSL/gcdata/geneseq/geneseq-emb1/AA1984.DAT:  
6: /SIDSL/gcdata/geneseq/geneseq-emb1/AA1985.DAT:  
7: /SIDSL/gcdata/geneseq/geneseq-emb1/AA1986.DAT:  
8: /SIDSL/gcdata/geneseq/geneseq-emb1/AA1987.DAT:  
9: /SIDSL/gcdata/geneseq/geneseq-emb1/AA1988.DAT:  
10: /SIDSL/gcdata/geneseq/geneseq-emb1/AA1989.DAT:  
11: /SIDSL/gcdata/geneseq/geneseq-emb1/AA1990.DAT:  
12: /SIDSL/gcdata/geneseq/geneseq-emb1/AA1991.DAT:  
13: /SIDSL/gcdata/geneseq/geneseq-emb1/AA1992.DAT:  
14: /SIDSL/gcdata/geneseq/geneseq-emb1/AA1993.DAT:  
15: /SIDSL/gcdata/geneseq/geneseq-emb1/AA1994.DAT:  
16: /SIDSL/gcdata/geneseq/geneseq-emb1/AA1995.DAT:  
17: /SIDSL/gcdata/geneseq/geneseq-emb1/AA1996.DAT:  
18: /SIDSL/gcdata/geneseq/geneseq-emb1/AA1997.DAT:  
19: /SIDSL/gcdata/geneseq/geneseq-emb1/AA1998.DAT:  
20: /SIDSL/gcdata/geneseq/geneseq-emb1/AA1999.DAT:  
21: /SIDSL/gcdata/geneseq/geneseq-emb1/AA2000.DAT:  
22: /SIDSL/gcdata/geneseq/geneseq-emb1/AA2001.DAT:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2605	100.0	704	22	AA831165
2	2605	100.0	977	22	AA831164
3	2605	100.0	1124	15	AA845440
4	2605	100.0	1124	16	AA873953
5	2605	100.0	1124	20	AA873938
6	2605	100.0	1124	22	AA865945
7	2407	92.4	1124	21	AA859046
8	2407	92.4	1124	21	AA859048
9	2407	92.4	1124	16	AA873951
10	2407	92.4	1124	16	AA873951
11	2393	91.9	1122	15	AA846627

12	2391.5	91.8	1117	15	AA853146
13	2352	90.3	1101	20	AA831160
14	865.5	33.2	1138	14	AA839820
15	851	32.7	1137	20	AA830319
16	844.5	32.4	1122	16	AA873954
17	844.5	32.4	736	22	AA851184
18	675.5	25.9	1094	14	AA839821
19	455	17.5	431	20	AA874444
20	327.5	12.6	434	22	AA866756
21	315	12.1	241	22	AA872712
22	315	12.1	384	22	AA875479
23	313.5	12.0	587	22	AA831826
24	313	12.0	878	22	AA838033
25	312.5	11.8	269	22	AA838031
26	307	11.6	497	22	AA866268
27	307	11.6	436	22	AA866269
28	305	11.7	474	22	AA872715
29	305	11.7	296	22	AA856069
30	305	11.7	299	21	AA876014
31	305	11.7	299	21	AA876077
32	305	11.7	299	22	AA855953
33	305	11.7	299	22	AA856016
34	300	11.5	310	22	AA872716
35	297	11.4	1050	22	AA866267
36	289.5	11.1	350	22	AA808075
37	286.5	11.0	594	22	AA858834
38	286	11.0	833	21	AA856116
39	286	11.0	833	21	AA860392
40	284.5	10.9	321	22	AA837639
41	277	10.6	129	22	AA832555
42	277	10.6	129	22	AA827637
43	277	10.6	866	22	AA860393
44	277	10.6	866	22	AA860394
45	272.5	10.5	456	22	AA832559

## ALIGNMENTS

RESULT 1

AA831165

1D AA831165 standard; protein; 704 AA.

XX

AC AA831165;

XX 02-APR-2001 (first entry)

XX Amino acid sequence of a Tek/Fc fusion protein.

DE Fusion protein; receptor tyrosine kinase; Tek; Fc portion;

KW immunoglobulin G1; IgG1, angiogenesis; tumour; cellular neovascularisation;

KW inflammatory disease; arthritis; rheumatism; psoriasis; eye disorder;

KW diabetic retinopathy; retinopathy; neovascular glaucoma; retinoblastoma;

KW retrolental fibroplasia; rubecosis; uveitis; macular degeneration;

KW graft neovascularisation; cancer; metastatic sarcoma; carcinoma;

KW wound granulation.

XX Synthetic.

OS Homo sapiens.

XX Key

FT Peptide

FT Location/Qualifiers

FT 1..18

FT /note= "signal peptide"

FT Protein

FT 19..472

FT /note= "Tek extracellular domain fragment"

FT Protein

FT 473..704

XX /note= "Fc portion"

XX W0200075323-A1.

XX 14-DEC 2000.

XX 07-JUN 2000, 2000WO-US15706.

PE

XX 07-JUN-1999; 990S-0137889.  
 XX (IMMV ) IMMUNEX CORP.  
 XX Carretti DP, Borges LG, Fanslow WC;  
 XX WPI; 2001-112149/12.  
 XX New Tek polypeptides antagonist having a fragment of the Tek  
 XX extracellular domain, useful for treating mammals with a disease  
 XX mediated by angiogenesis, e.g. tumors, ocular neovascularisation or  
 XX inflammatory diseases -  
 XX claim 17; Page 37-39; 43pp; English.  
 XX the present sequence represents a fusion protein comprising the  
 XX extracellular domain of the human receptor tyrosine kinase Tek and the  
 XX Fe portion of human immunoglobulin (Ig) G1. The Tek fragment lacks all  
 XX or part of the region containing fibronectin type III (FNIII) motifs, and  
 XX retains the ability to bind at least one Tek ligand. The fusion  
 XX polypeptide is an angiogenesis inhibitor, and a Tek antagonist. The  
 XX polypeptide or soluble Tek multimer, antibodies of antibody fragments are  
 XX useful for treating a mammal having a disease or condition mediated by  
 XX angiogenesis, e.g. a solid tumour a condition or disease characterized by  
 XX ocular neovascularisation. In particular, the Tek antagonists are useful  
 XX for treating or preventing inflammatory diseases (e.g. arthritis,  
 XX rheumatism or psoriasis), certain eye disorders (e.g. diabetic  
 XX retinopathy, retinopathy of prematurity, neovascular glaucoma,  
 XX retinoblastoma, retrolental fibroplasia, rubeosis, uveitis, macular  
 XX degeneration or graft neovascularisation), cancer (e.g. metastatic  
 XX sarcomas or carcinomas), or wound granulation.  
 XX Sequence 704 AA;  
 Query Match 100.0%; Score 2605; DB 22; Length 704;  
 Best Local Similarity 100.0%; Pred. No. 2,2e 187;  
 Matches 472; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MDLSASLVACGVSLISGIVEGAMDLIIINSIPLIVSDAFTSLITCIASGWRPHEPTICRD 60  
 DB 1 mdslaslvicqvsllsqivequmdlilinslplsdsdaetsllclclsgwrphcpilqlrd 60  
 QY 61 FEALMNOHQDPLLVTDVTRWAKKVVWKKPKASKINQAYVCHGVPCGEAIPETMKMPQ 120  
 DB 61 fealmnhdpllvtdvtrewakvkvwkkaskngayfcoegrvrggealirrtmkmrq 120  
 QY 121 QASFLPATLITVDRIDNVNITSPKKVLTKERDAVIYKKNSTHVSVPREVPDILEVHLPH 180  
 DB 121 qasflpatlmtvdrdkydrnntsfkkltkcedatlykngsfhsvprhcvpdiilevhlph 180  
 QY 181 AQWQACVVSARYICGNIFTSAPTHLIVHPCFAOKWGPCNHLITACNNNCVCHEDTQRC 240  
 DB 181 aqpdqavvsaryiqnlltsa:tlivrrceaqkwqecobletacmngvchedlqec 240  
 QY 241 QCHPQPCWRCACGAEHIFQKQKQKSCQCKSVYFCLADAYQCSA:QWKKQI:TW 300  
 DB 241 qchpmpcwrcacgahifqkqkqkscqcksvyfcldayqcsa:qwkq:tw 300  
 QY 301 ACHPFVYGPTEKLEYSNNKEMTDFEATGASPMWGLQAEFEKIPRMTPKIVLPLQHI 360  
 DB 301 achpfvygpteklefsnnkemtdfeatgaspwglqae:fe:iprmtpkivlplqhie 360  
 QY 361 VNSCKNPKICKASQWIPNFMETIVKQNGIVIHKINHNHNSVAFITTHRIILPPHSC 420  
 DB 361 vnsqknplckasqwpipneemtivkpdqivlhpknhnhdhsvalfithrilppdsq 420  
 QY 421 VWGCVNIVAGWVKPFNISVAIVPKPLNAPNVIHTGHNFANINISSPPVHG 472  
 DB 421 vwvsgvntvagnvkvpfnisvavpkplnapnvihtghnfahfainisspfyfg 472

RESULT 2  
 AAB31164  
 ID AAB31164 standard; protein; 977 AA.  
 XX  
 AC AAB31164;  
 XX  
 DT 02-APR-2001 (first entry)  
 XX  
 DE Amino acid sequence of a tek/tyc fusion protein.  
 XX  
 KW Fusion protein, receptor tyrosine kinase; Tek; Fe portion;  
 KW immunoglobulin G1; IgG1; angiogenesis; tumor; ocular neovascularisation;  
 KW inflammatory disease; arthritis; rheumatism; psoriasis; eye disorder;  
 KW diabetic retinopathy; retinopathy; neovascular glaucoma; retinoblastoma;  
 KW retrolental fibroplasia; rubeosis; uveitis; macular degeneration;  
 KW graft neovascularisation, cancer, metastatic sarcoma, carcinoma;  
 KW wound granulation.  
 XX  
 OS Synthetic.  
 OS Homo sapiens.  
 XX  
 FH Key location/Qualifiers  
 FT Peptide 1..18 "signal peptide"  
 FT Protein 19..745  
 FT Protein /note- "Tek extracellular domain"  
 FT Protein 746..977  
 FT Protein /note- "Fe portion"  
 XX  
 PN WC200075323-AL.  
 XX  
 PD 14-DEC-2000.  
 XX  
 XX 07-JUN-2000; 2000WO-US15704.  
 XX  
 XX 07-JUN-1999; 990S-0137889.  
 XX (IMMV ) IMMUNEX CORP.  
 XX  
 PI Carretti DP, Borges LG, Fanslow WC;  
 XX  
 DR WPI; 2001-112149/12.  
 XX  
 DR New Tek polypeptides antagonist having a fragment of the Tek  
 XX extracellular domain, useful for treating mammals with a disease  
 XX mediated by angiogenesis, e.g. tumors, ocular neovascularisation or  
 XX inflammatory diseases -  
 XX Claim 41; Page 34-37; 43pp; English.  
 XX The present sequence represents a fusion protein comprising the  
 XX extracellular domain of the human receptor tyrosine kinase Tek and the  
 XX Fe portion of human immunoglobulin (Ig) G1. The Tek fragment lacks all  
 XX or part of the region containing fibronectin type III (FNIII) motifs, and  
 XX retains the ability to bind at least one Tek ligand. The fusion  
 XX polypeptide is an angiogenesis inhibitor, and a Tek antagonist. The  
 XX polypeptide or soluble Tek multimer, antibodies or antibody fragments are  
 XX useful for treating a mammal having a disease or condition mediated by  
 XX angiogenesis, e.g. a solid tumor a condition or disease characterized by  
 XX ocular neovascularisation. In particular, the Tek antagonists are useful  
 XX for treating or preventing inflammatory diseases (e.g. arthritis,  
 XX rheumatism or psoriasis), certain eye disorders (e.g. diabetic  
 XX retinopathy, retinopathy of prematurity, neovascular glaucoma,  
 XX retinoblastoma, retrolental fibroplasia, rubeosis, uveitis, macular  
 XX degeneration or graft neovascularisation), cancer (e.g. metastatic  
 XX sarcomas or carcinomas), or wound granulation.  
 XX Sequence 977 AA;  
 Query Match 100.0%; Score 2605; DB 22; Length 977;  
 Best Local Similarity 100.0%; Pred. No. 4,2e 187;  
 Matches 472; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 MDSLASVLCGVSLLSSTVEGAMDLILINSLVSDAETSLTICASGWRPHEPITIGRD 60
DB 1 mdsiasvllcgvslisgstvegamlilinslvpsdaetslticagwprhepitigrd 60
QY 61 FEALMNOHQDPLEVTQDVTPEWAKKVVWKEKASKINGAYFCEGVPYCEPAIPTWKMQ 120
DB 61 fealmnohqdplevtqdvtpewakkkvwwkekaskingayficegrvgepaipr:mkmtq 120
QY 121 QASFLPATILMTVPGKGNVNISFKKVLKEEDAVIYKNGSFTHSVPRHEVPOILVHLPH 180
DB 121 qasflpatilmtvpgkgnvnisfkkvleedaviyknsgfthsvprhevpdilvhlph 180
QY 181 AQPGFARVYSAFYTGSLNFTSAFTGLVPR:EA:KWGPE:NHLCIA:MNGV:HEDTGEC 240
DB 181 aqpgfarysarytgnlftsaftglvpr:ea:kwgpe:nhlcia:mnngv:hedtgec 240
QY 241 ICPPGEMGRICEKAFELHIEGRT:KERTS:QEG:KSNVF:LPDPYQ:SC:ALGHWK:ILW:NE 300
DB 241 icppgemgricekafelhieg:rt:kerts:qeg:ksnvf:lpdpq:sc:alqwkglqne 300
QY 301 ACHPGFYGPDOCKLKCSCNNGEMDURFGCICSPQWQGLQCEKRGKLVHMTPKIVDLPHIE 360
DB 301 achpgfygpdcklkcscnngemodurfgcicspqwgqlqceqgiprmtpkivdlphie 360
QY 361 VNSGKFNPKICKASGWPLPTNEEMTLVPRDGTVLHPKDFNHTDHSVAIFTHIRLLPPDSG 420
DB 361 vnsqkfnpkickasgwplptneemtlvprdgvtlhpkdfnhtdhsvaifthrillppdsq 420
QY 421 VWVCSVNTVAGWVEKPFNISVKVLFKFLNAPNV:DTGHNFAVINISSEPYFG 472
DB 421 vwvcsvntvagwvekpfnis:kvlfkflnnpnv:dtghnfa:vinissepyfg 472

RESULT 3
ID AAR45440 standard; protein: 1124 AA.
XX AAR45440;
XX AC AAR45440;
XX DT 25-JUN-1994 (first entry)
XX DE Human orphan receptor kinase.
XX KW Ork; ligands; antibodies; PCR; amplification.
XX OS Homo sapiens.
XX FI Key Location/Qualifiers
FT Peptide 1..18
FT Region /note= "signal peptide"
FT Region 746..772
FT Region /note= "transmembrane region"
FT Region 211..340
FT Misc-difference 44 /note= "ECF-like repeat"
FT Misc-difference 44 /note= "Cys of immunoglobulin domain"
FT Misc-difference 102 /note= "Cys of immunoglobulin domain"
FT Misc-difference 19..1124 /note= "claimed fragment"
FT Misc-difference 19..745 /note= "claimed fragment"
XX WO9400469-A.
XX PD 06-JAN-1994.
XX PF 25-JUN-1993; 94W-050609
XX PR 26-JUN-1992; 92US 0905600.
XX PA (IMMUNEX ) IMMUNEX COPP

```

```

XX Ziegler SF;
XX WPI: 1994-026132/03.
XX N-PSDB; AAO55179.
XX DNA and protein sequences for orphan receptor tyrosine kinase -
XX and expression vectors for production of recombinant protein and
XX antibodies specific for the protein, useful in research
XX Claim 13; Fig 1; 57pp; English.
XX Degenerate oligonucleotide primers based on the sequence conserved
XX in the kinase domain of all receptor tyrosine kinases was used for
XX PCR of single stranded cDNA from human placental polyA mRNA. PCR
XX prod. HPK-6 contained a novel sequence which was used as a probe to
XX isolate longer fragments from a human placental cDNA library. One
XX clone contained the entire coding region and was called the ork gene.
XX The gene prod. shown can be used as a research tool in in vitro assays
XX for detection of ork, its ligands or their interactions.
XX See also AAR45441.
XX SQ Sequence 1124 AA;
XX Query Match 100.0%; Score 2605; DB 15; Length 1124;
XX Best Local Similarity 100.0%; Pred. NO. 3.8e-187;
XX Matches 472; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MDSLASVLCGVSLLSSTVEGAMDLILINSLVSDAETSLTICASGWRPHEPITIGRD 60
DB 1 mdsiasvllcgvslisgstvegamlilinslvpsdaetslticagwprhepitigrd 60
QY 61 FEALMNOHQDPLEVTQDVTPEWAKKVVWKEKASKINGAYFCEGVPYCEPAIPTWKMQ 120
DB 61 fealmnohqdplevtqdvtpewakkkvwwkekaskingayficegrvgepaipr:mkmtq 120
QY 121 QASFLPATILMTVPGKGNVNISFKKVLKEEDAVIYKNGSFTHSVPRHEVPOILVHLPH 180
DB 121 qasflpatilmtvpgkgnvnisfkkvleedaviyknsgfthsvprhevpdilvhlph 180
QY 181 AQPGFARVYSAFYTGSLNFTSAFTGLVPR:EA:KWGPE:NHLCIA:MNGV:HEDTGEC 240
DB 181 aqpgfarysarytgnlftsaftglvpr:ea:kwgpe:nhlcia:mnngv:hedtgec 240
QY 241 ICPPGEMGRICEKAFELHIEGRT:KERTS:QEG:KSNVF:LPDPYQ:SC:ALGHWK:ILW:NE 300
DB 241 icppgemgricekafelhieg:rt:kerts:qeg:ksnvf:lpdpq:sc:alqwkglqne 300
QY 301 ACHPGFYGPDOCKLKCSCNNGEMDURFGCICSPQWQGLQCEKRGKLVHMTPKIVDLPHIE 360
DB 301 achpgfygpdcklkcscnngemodurfgcicspqwgqlqceqgiprmtpkivdlphie 360
QY 361 VNSGKFNPKICKASGWPLPTNEEMTLVPRDGTVLHPKDFNHTDHSVAIFTHIRLLPPDSG 420
DB 361 vnsqkfnpkickasgwplptneemtlvprdgvtlhpkdfnhtdhsvaifthrillppdsq 420
QY 421 VWVCSVNTVAGWVEKPFNISVKVLFKFLNAPNV:DTGHNFAVINISSEPYFG 472
DB 421 vwvcsvntvagwvekpfnis:kvlfkflnnpnv:dtghnfa:vinissepyfg 472

RESULT 4
AAR73953
ID AAR73953 standard; Protein; 1124 AA.
XX AAR73953;
XX AC AAR73953;
XX DT 22-JAN-1996 (first entry)
XX DE Human TRK tyrosine kinase protein.
XX KW tie-2; receptor-tyrosine kinase; DNA primer; cancer; angiogenesis;

```

KW vasculogenesis: tek.

OS Homo sapiens.

PN WC9513387-A1.

XX 18-MAY-1995.

XX 12-NOV-1994; 94WO-EU03767.

XX 12-NOV-1993; 93US-0152552.

PA (PLAC ) MAX PLANCK GES FORDERUNG WISSENSCHAFTEN.

PI Risau W;

DR WP1; 1995-194105/25.

XX New tie-2 receptor tyrosine kinase and related nucleic acid and  
PT methods for detecting tie-2 modulators for treating eq cancer,  
PI associated with angiogenesis and vasculogenesis

PS Disclousure; Page 49; 81pp; English.

XX This protein is the human homolog of mouse tie-2 receptor  
CC tyrosine kinase.

XX Sequence 1124 AA;

Query Match 100.0%; Score 2605; DB 16; Length 1124;

Best Local Similarity 100.0%; Prod. No. 3.8e-187;

Matches 472; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDSLASIVCGVSLISGIVGAMDLILINSPLVSDAETSLTICASQWRPHEPTIGRD 60

Db 1 mdsaslvicgvsllsgtvegamdlilinsplvsdaetslticassqwrpheapitard 60

QY 61 FTALMNGQDPLLEVTDVTRWAKVWKKRASKINGAYFCGKVGCEAIRITMKMRQ 120

Db 61 ftalnmghdpllevtdvtrwawkkvwwkrekackindayfcdgvrfgcairirtmkmrq 120

QY 121 QASFLPATITMTVDKGNVNI SFKKVLTKEDDAVIYKNGSPFHSVPRFVDPDILEVHLPH 180

Db 121 qasflpatirtmvdkgndvni sfkvvtkedavtykngsfhsvprfhpdpdilevhlph 180

QY 181 AQPQDAGVYSARYTGNLTSAFTPLIVPPGCEAQRWGPENHILATAMNNVWVHEETGEC 240

Db 181 aqpdaqvysarytgnltsaftplivrrceaqkwpqecnhlctacmnqqvchedtqec 240

QY 241 ICNPGMGRNCEKACELHTFGTCKRQSGQPGCKSVYFCLLHPYQSCAIGWKGQCNF 300

Db 241 icpplmgtrekeacelhtfgtckrqsqpgcksvyfccllhpypqscalgwkgqcnf 300

QY 301 ACHPFPYQDCKRQSRNKGEMFTSTQSTSPQWNRNLFEPFIPMTKPIVGLPHITF 360

Db 301 achpfpyqdkrqsrnkgemfstdstspqwnrnlfeppfipmtkpiivglphtf 360

QY 361 VASQKFNPKASQWPIFNELMVLVHSGVLVHRKPEFNHILATAMNNVWVHEETGEC 420

Db 361 vasqkfnpkasqwpifnlemlvlvhsqvlvhrkpefnhiltatamnnvwvheetg 420

QY 421 VWVGSVNTVACWVERFENLVKVLKPLNATVNCVCHNFAVNISSLPVFC 472

Db 421 vwvgsvntvawverfenlvkvlkplnajtncvchnfavnisslpvfc 472

RESULT 5

AAV30318

XX AAY30318 standard; protein; 1124 AA.

XX

AC AAY30318;

XX

DI 15-NOV-1999 (first entry)

XX Amino acid sequence of TEK (also known as TIE2) protein.

XX TEK protein; TIE2 protein; receptor tyrosine kinase; T cell response;

KW immune response; endothelial cell; tumor-associated vasculature;

XX coagulation; thrombosis; cancer; anticancer vaccine.

OS Homo sapiens.

PN WO9943801-A1.

XX 02-SEP-1999.

XX 26-FEB-1999; 99WO-GB00583.

XX 26-FEB-1998; 96CH-0004121.

XX (CANC-) CANCER RES CAMPAIGN TECHNOLOGY.

XX Durrant LG, Hewett PW, Ramage JM, Spendlove I;

XX WP1; 1999-540586/45.

XX New peptides containing at least one epitope from tek receptor  
CC tyrosine kinase, used in vaccines against cancer

XX Disclousure; Fig 1; 56pp; English.

XX The present sequence represents the TEK protein, also known as TIE2  
CC protein. TEK is a receptor tyrosine kinase. TEK contains epitopes  
CC which bind to MHC. The presentation of TEK epitopes can also stimulate  
CC helper cell and/or cytotoxic T cell responses. The immune response is  
CC directed against endothelial cells in the tumor-associated  
CC vasculature and includes production of antibodies that bind to the  
CC cells, causing coagulation and thrombosis. The immune response is  
CC targeted to endothelial cells lining blood vessels of the tumor (these  
CC cells overexpress tek). so damage to even a few cells will kill many  
CC tumor cells. These target cells are accessible to the immune response  
CC and problems of antigenic heterogeneity, MHC loss and resistance to  
CC apoptosis (associated with epithelial cells) are unlikely to occur in  
CC normal endothelial cells. TEK epitopes (see AAY30320-24) are used to  
CC generate antibodies and for prevention and treatment of cancer.  
CC The peptides, and recombinant DNA constructs or viral vectors that  
CC express them, are useful as anticancer vaccines to target endothelial  
CC cells that line blood vessels of the tumor.

XX Sequence 1124 AA;

Query Match 100.0%; Score 2605; DB 20; Length 1124;

Best Local Similarity 100.0%; Prod. No. 3.8e-187;

Matches 472; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDSLASIVGVSLISGIVGAMDLILINSPLVSDAETSLTICASQWRPHEPTIGRD 60

Db 1 mdsaslvicgvsllsgtvegamdlilinsplvsdaetslticassqwrpheapitard 60

QY 61 FTALMNGQDPLLEVTDVTRWAKVWKKRASKINGAYFCGKVGCEAIRITMKMRQ 120

Db 61 ftalnmghdpllevtdvtrwawkkvwwkrekackindayfcdgvrfgcairirtmkmrq 120

QY 121 QASFLPATITMTVDKGNVNI SFKKVLTKEDDAVIYKNGSPFHSVPRFVDPDILEVHLPH 180

Db 121 qasflpatirtmvdkgndvni sfkvvtkedavtykngsfhsvprfhpdpdilevhlph 180

QY 181 AQPQDAGVYSARYTGNLTSAFTPLIVPPGCEAQRWGPENHILATAMNNVWVHEETGEC 240

Db 181 aqpdaqvysarytgnltsaftplivrrceaqkwpqecnhlctacmnqqvchedtqec 240

QY 241 ICNPGMGRNCEKACELHTFGTCKRQSGQPGCKSVYFCLLHPYQSCAIGWKGQCNF 300

Db 241 icpplmgtrekeacelhtfgtckrqsqpgcksvyfccllhpypqscalgwkgqcnf 300

QY 301 ACHPEFYSPGCKKASNNEMTUPPAPVLTSPWQSLPEPSIPAMTPKIVDLPDITE 460  
 DB 301 achpfyfgpckkrescunymedrfgcletspgswgglqeragipmtpkivdlpdhie 360

QY 361 VNSCKENPICKASWPLTNEMTLVKPDCTVTHPRDNFNIHTHFSVAITTHIRLPPNSC 420  
 DB 361 vnscknpickaswplneemtlvkdpcvtvthprdnfnithfsvaitthirllppdsq 420

QY 421 VVCSVNTVAGMVKPFNISVKVLPKPLNAPNVIDTCHNPAVINISSEYVFG 472  
 DB 421 vvcsvntvagmvekpnisvkvlpkplnapnvidtchnpavinissepyfq 472

RESULT 6  
 AAG65945  
 ID AAG65945 standard; protein: 1124 AA.  
 AC AAG65945;  
 XX  
 DT 11-FEB-2002 (first entry)  
 DE Amino acid sequence of human Tie-2.  
 XX  
 KW Tie-2, catalytic domain, receptor tyrosine kinase, crystalline; human;  
 KW cytosolic; vasotropic; anti-anemic; antiarteriosclerosis; nephrotropic;  
 KW ophthalmological; hepatotropic; antithyroid; antiinflammatory; antidiabetic;  
 KW gastrointestinal; antirheumatic; osteopathic; antiarthritic; hemostatic;  
 KW antipsoriatic; dermatological; immunosuppressive; antibacterial  
 OS Homo sapiens  
 XX  
 FH Key Location/Qualifiers  
 FT Domain 802..1124  
 FT /note="catalytic domain, specifically claimed fragment"  
 XX  
 PN WC2000172778-A2.  
 PD 04-OCT-2001.  
 XX  
 PF 20 MAR 2001, 2001WO 0538833.  
 XX  
 PP 29-MAR 2000; 2000US-192920P  
 PA (BAP) BASF AG.  
 XX  
 PI Rump NJ, Arnold LD, Dixon RW, Hoeffken HW, Ailen K, Bellamacina C;  
 DR WPI; 2001-648437/74.  
 XX  
 PT Crystalline polypeptide useful for identifying inhibitors of a Tie-2  
 PT protein as well as determining the three dimensional structure of a  
 PT polypeptide comprising the catalytic domain of a Tie-2 polypeptide -  
 XX  
 PS Claim 6; Fig 1; 242pp; English.  
 CC The invention relates to a crystalline polypeptide, comprising the  
 CC catalytic domain of a receptor tyrosine kinase Tie-2 protein. The  
 CC crystalline forms are useful for identifying inhibitors of a Tie-2  
 CC protein as well as determining the three dimensional structure of a  
 CC the catalytic domain of a Tie-2 polypeptide. A Tie-2 inhibitor may be  
 CC used to treat a Tie-2 dependent condition in a patient (especially a  
 CC human), where the condition is characterized by excessive vascular  
 CC proliferation e.g. a hyperproliferative disorder, cancer (e.g. sarcoma,  
 CC osteoma, melanoma, lymphoma, and leukemia), a cardiovascular condition  
 CC (e.g. atherosclerosis, ischemia, anemia, and vascular leakage disorders),  
 CC an ocular condition (myopia, chronic retinal detachment, conjunctivitis,  
 CC retinopathy, and macular degeneration), von Hippel Lindau disease,  
 CC pemphigoid, psoriasis, Paget's disease, polycystic kidney disease,  
 CC fibrosis, sarcoidosis, cirrhosis, thyroiditis, Osher-Weber-Rendu disease,  
 CC chronic inflammation, synovitis, inflammatory bowel disease, Crohn's  
 CC disease, rheumatoid arthritis, osteoarthritis, psoriatic arthritis, and  
 CC ulcer or sepsis, especially where the disorder involves aberrant

CC endothelial-periendothelial interactions. The Tie-2 inhibitor may be  
 CC used to decrease fertility, and promote angiogenesis or vasculogenesis  
 CC (in combination with a pro-angiogenic growth factor). The present  
 CC sequence represents the human tie 2 protein sequence.  
 XX  
 SQ Sequence 1124 AA;  
 Query Match 100.0%; Score 2605; DB 22; Length 1124;  
 Best local Similarity 100.0%, Pred. No. 3.8e-187;  
 Matches 472, Conservative 0, Mismatches 0, Indels 0, Gaps 0;  
 QY 1 MDSLASLVLCVSVLLISSTVEGAMDLILINSLPLVSDAETSLTCTASQWPPPIPTITQD 60  
 DB 1 mdslasvlcvsvllisgtvegamdliilnslplvsdaetsltctasgwppphtitgrd 60  
 QY 61 FEALMNOHQHPIFVTQVTPFWAKKVVWKKFKASKINGAYFCKGPGVGGAIPIPTMKMQ 120  
 DB 61 fealmnohqhpiqvtpfwakkvvwkkfkaskingayfckgpgvggairptmkmq 120  
 QY 121 QASFLPATLIMTVKGNVNISPKKVLIKKEDAVIYKNGSFIHSPRHHVPPDILRVHDPH 180  
 DB 121 qasflpatlmtvdkgnvnisppkvlirkedaviykngsfihsprrhhvppdilrvhph 180  
 QY 241 LPPQIMARCEKACELHIGRLCKERCSCQEQAKSYVFLPDYGSALQWKLQCNH 300  
 DB 241 lppqimartceakcelhigrckercscqeqaksvyflpdpygsalqwkqlqne 300  
 QY 361 ACHPEFYSPGCKKASNNEMTUPPAPVLTSPWQSLPEPSIPAMTPKIVDLPDITE 360  
 DB 361 achpfyfgpckkrescunymedrfgcletspgswgglqeragipmtpkivdlpdhie 360  
 QY 421 VVCSVNTVAGMVKPFNISVKVLPKPLNAPNVIDTCHNPAVINISSEYVFG 472  
 DB 421 vvcsvntvagmvekpnisvkvlpkplnapnvidtchnpavinissepyfq 472

RESULT 7  
 AAY59046  
 ID AAY59046 standard; protein: 1118 AA.  
 AC AAY59046;  
 XX  
 DT 07-MAR-2000 (first entry)  
 DE Mouse receptor tyrosine kinase protein.  
 XX  
 KW Receptor tyrosine kinase protein; angiogenesis; cardiogenesis; mouse;  
 KW tumorigenesis.  
 XX  
 OS Mus musculus.  
 PN M55498187-A  
 XX  
 PD 07-DEC-1999.  
 XX  
 PF 23-APR-1997; 97US-0838957.  
 XX  
 PP 20-MIL-1404; 94US-0278089.  
 PP 40-ILL-1992; 92US-0021795.  
 PP 29-APR-1994; 94US-0235408.  
 XX  
 PA (MOUN) MOUNT SINAI HOSPITAL CORP.  
 XX Yamauchi TP, Breitman T, Dumont DT, Passant J, Breitman ML;  
 PI

```

XX DR WPI: 2000-052545/04.
XX N PSDB: AA240638.
XX PT Novel receptor tyrosine kinase protein involved in angiogenesis.
XX PT cardiotogenesis and tumorigenesis -
XX PS Claim 4; Fig 1; 123pp; English.
XX CC The invention provides novel receptor tyrosine kinase protein sequences
XX CC (AA59046-48) and nucleic acids encoding the polypeptides. The isolated
XX CC and purified polypeptides are useful for studying the developmental
XX CC expression of a receptor tyrosine kinase protein which may have a role
XX CC in angiogenesis, cardiotogenesis and tumorigenesis. The present sequence
XX CC represents a mouse receptor tyrosine kinase protein.
XX PT Sequence 1118 AA;
XX PS
XX CC Query Match: 92.4%; Score 2407; DB 21; Length 1118;
XX CC Best Local Similarity: 91.9%; Pred. No. 2.8e-172;
XX CC Matches 434; Conservative 17; Mismatches 21; Indels 0; Gaps 0;
QY 1 MDSLASLVLCGVSLLSGTVEGAMDILLINSLPLVSDAETSLTCLASGWRPHEPITGRD 60
Db 1 mdsfaglv cgvsllygvvsgamdillinsplvsdaetsltclasswplgtpitgrd 60
QY 61 FEALMNHQDPLEVTQDVTREWAKKVVWKKREKASKINAYFEGFVPPVEALPVPIMKMQ 120
Db 61 fealmnhqdpdvtdvtrewakkvwkkrekaskinayfegfvppvealpvipimkmrq 120
QY 121 QASFLPATLTMVVDKGVNVSFPPKVLKKEADAVLYKNGSPFHSVPREHVPDILEVHLPH 180
Db 121 qasflpatltmvdkgvnvsfppkvllkheadavlyknsgpfhsvprehvpdilevhlph 180
QY 181 AQPQDAGVYSARYIGNLTSAFTPIVPPCFAPKFWPEPCNHLTACMNNGVVEDTCEG 240
Db 181 aqpqdagvysaryignltsaftpiivppcfapkwfepcnnhlta cmnngvvedtceg 240
QY 241 ICPPGFMGRTECAKAELEHIFGHTCKEKSSQEGGKSYVFLPDPYGCSCATGKGLQNE 300
Db 241 icppgfmgrtceakaelelhifghtckekssqeggk syvflpdp ygcs catgkglqne 300
QY 301 ACHPGFYGHVCKI PSCNNGNFWCNEFGGCLCSQCGGLQCHPEGIPRMTPKIVLDPHLE 360
Db 301 achpgfyghvcki pscnngnfwcnefggclcsqcgglqchpegiprmtpkivldphle 360
QY 361 VNSGKFNPKCKASGWPLPTNEEMTLVKPDGTVLHPKDFNHIDHFSVAIFTIHRILPPDSG 420
Db 361 vnschkfnpkckasgwplptneemtlvkpdgtvlhpkdfnhidhfsvalftihrilppdsq 420
QY 421 VWVCVNTVAGVVEKPFNLSVKKVLPKPLNAPRVLTGHNFAVINISSEPYFG 472
Db 421 vwvcvntvagvvekpfnlsvkvlpkplnaprvltghnfvainisscpyfg 472
RESULT 8
AA59048
ID AA59048 standard; Protein: 1123 AA.
XX AC
XX AC AA59048;
XX AC
XX AC 07 MAR 2000 (first entry)
XX DT
XX DE Mouse tek receptor tyrosine kinase protein
XX KW Receptor tyrosine kinase protein, angiogenesis, cardiotogenesis; mouse;
XX KW tumorigenesis.
XX OS Mus musculus.
XX XX
XX IN US5998187-A.
XX XX

```

```

PD 07-DEC-1999.
XX XX
XX PF 23-APR-1997; 97US-0838957.
XX XX
XX PP 20-MIL-1994; 94US-0278089.
XX PP 30-JUL-1992; 92US-0421795.
XX PR 29-APR-1994; 94US-0235408.
XX XX
XX PA (MOJN ) MOUNT SINAI HOSPITAL CORP.
XX XX
XX PI Yamauchi TP, Breitman J, Dumont EJ, Rossant J, Breitman ML;
XX XX
XX PP WPI: 2000-052545/04.
XX DR N PSDB: AA240640.
XX XX
XX PT Novel receptor tyrosine kinase protein involved in angiogenesis.
XX PT cardiotogenesis and tumorigenesis -
XX PS Claim 1; Fig 11B; 123pp; English.
XX CC The invention provides novel receptor tyrosine kinase protein sequences
XX CC (AA59046-48) and nucleic acids encoding the polypeptides. The isolated
XX CC and purified polypeptides are useful for studying the developmental
XX CC expression of a receptor tyrosine kinase protein which may have a role
XX CC in angiogenesis, cardiotogenesis and tumorigenesis. The present sequence
XX CC represents a mouse tek receptor tyrosine kinase protein.
XX PS Sequence 1122 AA;
XX CC
XX CC Query Match: 92.4%; Score 2407; DB 21; Length 1122;
XX CC Best Local Similarity: 91.9%; Pred. No. 2.8e-172;
XX CC Matches 434; Conservative 17; Mismatches 21; Indels 0; Gaps 0;
QY 1 MDSLASLVLCGVSLLSGTVEGAMDILLINSLPLVSDAETSLTCLASGWRPHEPITGRD 60
Db 1 mdsfaglv cgvsllygvvsgamdillinsplvsdaetsltclasswplgtpitgrd 60
QY 61 FEALMNHQDPLEVTQDVTREWAKKVVWKKREKASKINAYFEGFVPPVEALPVPIMKMQ 120
Db 61 fealmnhqdpdvtdvtrewakkvwkkrekaskinayfegfvppvealpvipimkmrq 120
QY 121 QASFLPATLTMVVDKGVNVSFPPKVLKKEADAVLYKNGSPFHSVPREHVPDILEVHLPH 180
Db 121 qasflpatltmvdkgvnvsfppkvllkheadavlyknsgpfhsvprehvpdilevhlph 180
QY 181 AQPQDAGVYSARYIGNLTSAFTPIVPPCFAPKFWPEPCNHLTACMNNGVVEDTCEG 240
Db 181 aqpqdagvysaryignltsaftpiivppcfapkwfepcnnhlta cmnngvvedtceg 240
QY 241 ICPPGFMGRTECAKAELEHIFGHTCKEKSSQEGGKSYVFLPDPYGCSCATGKGLQNE 300
Db 241 icppgfmgrtceakaelelhifghtckekssqeggk syvflpdp ygcs catgkglqne 300
QY 301 ACHPGFYGHVCKI PSCNNGNFWCNEFGGCLCSQCGGLQCHPEGIPRMTPKIVLDPHLE 360
Db 301 achpgfyghvcki pscnngnfwcnefggclcsqcgglqchpegiprmtpkivldphle 360
QY 361 VNSGKFNPKCKASGWPLPTNEEMTLVKPDGTVLHPKDFNHIDHFSVAIFTIHRILPPDSG 420
Db 361 vnschkfnpkckasgwplptneemtlvkpdgtvlhpkdfnhidhfsvalftihrilppdsq 420
QY 421 VWVCVNTVAGVVEKPFNLSVKKVLPKPLNAPRVLTGHNFAVINISSEPYFG 472
Db 421 vwvcvntvagvvekpfnlsvkvlpkplnaprvltghnfvainisscpyfg 472
RESULT 9
AA59051
ID AA59051 standard; Protein: 1123 AA.
XX AC
XX AC AA59051;
XX XX

```

DT	21-JAN-1996	(first entry)
DE	Mouse tie-2 receptor tyrosine kinase protein.	
DE		
KW	tie-2; receptor-tyrosine kinase; DNA primer, cancer, angiogenesis,	
KX	vasculogenesis.	
KW		
KX		
OS	Mus musculus.	
XX		
Key	Location/Qualifiers	
FT	Cleavage-site 22..23	
FT	/note= "signal peptide cleavage site"	
FT		
FT	Misc-difference 44	
FT	/note= "involved in sulphydryl bonding of Ig domain"	
FT		
FT	Misc-difference 102	
FT	/note= "involved in sulphydryl bonding of Ig domain"	
FT	Region 211..251	
FT	/note= "epidermal growth factor-like repeat"	
FT	Region 255..298	
FT	/note= "epidermal growth factor-like repeat"	
FT	Region 302..340	
FT	/note= "epidermal growth factor-like repeat"	
FT	Domain 437..720	
FT	/note= "includes 3 fibronectin type-III domains"	
FT	Domain 829..921	
FT	/note= "tyrosine-kinase domain"	
FT	Domain 936..1081	
FT	/note= "tyrosine-kinase domain"	
FT	Region 135..137	
FT	/note= "KCD triplet"	
XX		
WO	W04513387-A1.	
PN	18-MAY-1995.	
PD		
PP	12-NOV-1994; 94WO-EP03767.	
PP		
PR	12-NOV-1993; 93OS-0152552.	
PA	(PLAC ) MAX PLANCK GES FÖRERPEUNG WISSENSCHAFTEN.	
PB	Risau W;	
PI		
XX		
DR	WPI: 1995-194105/25.	
DR	N-PSDB: AAQ91999.	
PT		
PT	New tie-2 receptor tyrosine kinase and related nucleic acid - and	
PT	methods for detecting tie 2 modulators for treating eg cancer,	
PT	associated with angiogenesis and vasculogenesis	
XX		
PS	Claim 3: Page 45; 81pp: English.	
XX		
CC	This protein may be expressed recombinantly in a host cell.	
CC	The expressed protein may be used to treat diseases, or	
CC	processes associated with angiogenesis and vasculogenesis, or	
CC	cancer. Cells that express the protein are used in screening	
CC	procedures and recombinant protein can be used for affinity	
CC	purification of tie-2 ligand.	
XX		
SQ	Sequence 1123 AA:	
	Query Match: 92.4%; Score 2407; DB 16; Length 1123;	
	Best Local Similarity 91.9%; Fred No. 2.8e+172;	
	Matches 434; Conservative 17; Mismatches 21; Indels 0; Gaps	
OY	1 MDLSLASLVAGCVSLISGVIRGAMDLIIINSIPVVSAPTSITCIASGWPHPEPTICRD 60	
Dd	1 mdslasvlcvsvlllygvvqamdiilinslpvsdaetllciagwhphpeptlgrrd 60	
OY	61 FEALMNHQDPLEVTQDVTEHNAKKVVKREKASKINGATPCGGVRGEAIRIRTMMKMQ 120	
Dd	61 fealmnhgdpdvtehvtehwkvwvkrekaskindayfcedrvfgqairirtmmkra 120	





FT Region 840...857  
 FT /label= conserved\_region\_II  
 FT 824...829  
 FT /note= "GXGXXG motif,  
 FT part of consensus ATP-binding site"  
 FT Binding-site 846...848  
 FT /note= "AXK motif,  
 FT part of consensus ATP-binding site"  
 FT Region 858...873  
 FT /label= conserved\_region\_III  
 FT 874...888  
 FT /label= conserved\_region\_IV  
 FT 889...907  
 FT /label= conserved\_region\_V  
 FT 908...928  
 FT /label= insert  
 FT 929...969  
 FT /label= conserved\_region\_VI  
 FT 970...988  
 FT /label= conserved\_region\_VII  
 FT 975...977  
 FT /note= "motif found in almost all known kinases"  
 FT 989...1003  
 FT /label= conserved\_region\_VIII  
 FT 997...1003  
 FT /note= "motif found in transmembrane RTKs"  
 FT 1004...1034  
 FT /label= conserved\_region\_IX  
 FT 1035...1055  
 FT /label= conserved\_region\_X  
 FT /note= "the tek amino acid sequence appears  
 FT e somewhere in the specification with  
 FT the residues Et at positions 1034-1040  
 FT rather than DV"  
 FT Region 1056...1085  
 FT /label= conserved\_region\_XI  
 FT 1086...1117  
 FT /label= conserved\_region\_XII  
 XX CA2085291-A.  
 XX 31-JAN-1994.  
 XX 14-DEC-1992; 92CA-2085291.  
 XX 30-JUL-1992; 92US-0921795.  
 XX (MOUN ) MOUNT SINA; HOSPITAL CORP.  
 XX Breitman ML, Dumont DJ, Rossant J, Yamaguchi TP;  
 XX WPI: 1994-126948/16.  
 XX N-PSDB: AAQ62129.  
 PT Receptor tyrosine kinase isolated from cells of the endothelial  
 PT lineage - used for identifying (ant)agonists of ligand-receptor  
 PT binding  
 XX Claim 9: Fig 1: 51pp; English.  
 XX RNA from 9.5 and 12.5 day old mice was used to synthesize cDNA by RT  
 CC PCR using primers previously demonstrated to preferentially amplify  
 CC tyrosine kinase sequences. Clones corresp. to 5 distinct tyrosine  
 CC kinases expressed during murine cardiogenesis were identified. Four  
 CC were known tyrosine kinases but one represented a new tyrosine kinase  
 CC which was designated "tek". AAK53146 is the deduced tek amino acid  
 CC sequence. Comparisons to other tyrosine kinases showed that the  
 CC putative kinase domain contains several conserved sequence motifs  
 CC (see features table).  
 XX Sequence 1117 AA:  
 SQ

Query Match 91.8%; Score 2391.5; DR 15; Length 1117;  
 Best local Similarity 91.7%; Pred. No. 4,1e-171;  
 Matches 433; Conserved 17; Mismatches 21; Indels 1; Gaps 1;  
 QY 1 MDSLASLVLCGVSLILSGTVEGAMDILINSLPLVSDAETSLTCTASGWPPEPTIGRD 60  
 DB 1 mdsilagivlcgvsililgyvvgamdilinsiplvsdaetsltctasgwppeptigrd 60  
 QY 61 FEALMNGHQPPLFVTOVTPFWAKKVMWPEKASKINGAYFGGSGVPGFAIPIPTMKMPQ 120  
 DB 61 fealmnqhqpplfvtovtptfwakkvmwpekaskinqayfegsgvpgfaipiptmkmrq 120  
 QY 121 QASPLPATLIMTVKGNVN:SPKKVLIKHDAVIVKNSPIHNSVPRHVPDILKVLPH 180  
 DB 121 qasplpatlilmvkgvnvnsppkkvlikhdavivknspihnsvprhvpdilkvlph 180  
 QY 181 ACPALGVYSAPVIGNLFTSAFTPIVPEPEALPWTFEINILGLIA'MNNEW'HEUTDEC 240  
 DB 181 acpalgvysapvignlftsaftrivpepealpwtfenilglia'mnnew'heutdec 240  
 QY 241 ACPALGVYSAPVIGNLFTSAFTPIVPEPEALPWTFEINILGLIA'MNNEW'HEUTDEC 240  
 DB 241 acpalgvysapvignlftsaftrivpepealpwtfenilglia'mnnew'heutdec 240  
 QY 301 ACHGCFYGPUCCKLKCSCNNGEMCDKQGGCLCSPLWGLQCEKKGIPKMIPIKIVDLPDHE 360  
 DB 300 acpsqyygpcklrchctneelcdriqgcscsqwqlqcekgprmpipqiedipdhie 359  
 QY 361 VNSGFNPICKASQWPIPTNFMETLVKPDQTVLIRPDNFNITDFSVATITHTPTTPPSG 420  
 DB 360 vnskfnpickasywpiptseentlvkpdutvlirpdnfytrfsvalftvrvlppdsq 419  
 QY 421 VWVSVNTVAGMVEKPFNISVKVLKPIINAPVINTGHNFVNTISSEPYFG 472  
 DB 420 vwsvntvagmvekfpinisvkvlpeptkapnvidghnfainassepyfg 471  
 RESULT 13  
 AAW83160  
 ID AAW83160 standard; Protein; 1101 AA.  
 XX  
 AC AAW83160;  
 XX  
 DT 11-FEB-1999 (first entry)  
 XX  
 DE Rat orphan tyrosine kinase receptor protein Rtk 7 (Tie 2).  
 XX  
 KW Receptor tyrosine kinase, Ror 1, Ror 2, Etk-1, Etk-2, detection,  
 KW neurotrophin activity; trkB; proto-oncogene; tyrosine kinase receptor;  
 KW binding protein; btk; Nf-3; diagnosis.  
 XX  
 OS Rattus sp.  
 XX  
 PN US5843749-A.  
 XX  
 PD 01-DEC-1998.  
 XX  
 PF 06-JUN-1995; 95US-0469537.  
 XX  
 PR 17-MAR-1995; 95US-0406247.  
 PR 26-JUL-1991; 91US-0736559.  
 PR 28-OCT-1993; 93US-0144992.  
 PR 06-JUN-1995; 95US-0469537.  
 XX  
 PA (FENSE-) REGENERON PHARM INC  
 XX  
 PI Maisongier PC, Maslakowski P, Yancopoulos GB;  
 XX WPI: 1999-044584/04.  
 XX  
 PT DNA encoding receptor tyrosine kinase proteins - and corresponding  
 PT proteins  
 XX

Example: Fig 20; 194pp; Emailsh.

PS the present invention describes nucleic acid molecules for ror-1,  
 CC ror-2, chk 1 and chk 2. Also described are the corresponding proteins;  
 CC ror-1, ror 2, chk 1, and chk 2. the proteins are orphan tyrosine  
 CC tyrosine kinases. The present sequence represents rat orphan tyrosine  
 CC kinase receptor Rtk-7 (Tie-2) from the present invention.  
 XX  
 SQ Sequence 110; AA;

Query Match 90.3%; Score 2352; Db 20; Length 1101;

Best Local Similarity 93.4%; Pred No 3 8e-168;

Matches 423; Conservative 12; Mismatches 18; Indels 3; Gaps 0;

QY 20 VEGAMDLILINSPLVSVDAFTSLTCLASQWRHPEPTIGRDPEALMNOHQDPLEVTDVDT 79

Db 1 veyamdliinsplvsdaftsltlcasqwhphtlqrdlealmnqdpplvtdvdt 60

QY 80 REWAKKVMKREKAKINAYFEGEVRGEATPTPTMKPQCAFLPATLMTVDRGDNV 139

Db 61 rewakkvwmrkaskingayfegvrgga:rtmkmgqasfipatltmtvdrdgv 120

QY 140 NISFKKVLKEEDAVYKNGSFHHSVPREVPDILEVILPHAQPDAGVYSARYIGNLF 199

Db 121 nstfkvlkiedavikngsflhsvprhevpdlilevhlphapqdaqvysaryignlf 180

QY 200 ISAPTRRLVRRCAKWKIPENACIAACNNQVCHEDIQEGICHPGPMGRICCAACELHT 259

Db 181 tsatrlivrrcaqkwipdn:cpctickngsvchedlqecicppalmgrtceakcephl 240

QY 260 FETPTKFPKSGEYKSVYFTPTGYVSSATLWKGLQNEAHPGYSYPLKLEWLCNN 319

Db 241 ftrtkeressgeykyvftfpgyssatfwglp:nea:fygylp:ekl:cheth 300

QY 320 GEMCDREFOALCSQWQWQTLQEPETIPMTKPIVLPDHEVNSKFNPLCKASGWLPT 379

Db 301 eemcdrlqecleswqlqlq:ekarpmt:qjledlph:evnsqklupickasgwplpt 360

QY 380 NEMHILVKVIGTIVLHKIPNHDHFSVAFTIHRILPDSQWVWCSVNTVAGVVEKPFNI 439

Db 361 seemlvkpdqtlvlpadlnhldhsvaitvnrilpdsq:wvcsvntv:agmvekpini 420

QY 440 SVKVLKPKLPANVITDITGNEFAVINISSEPYEG 472

Db 421 svkvlkpkp:thefevtditg:nefa:vinissepyfg 453

RESULT 14

AAR39820

XX AAR39820 standard; Protein: 1138 AA.

XX AAR39820;

XX 12 JAN 1994 (first entry)

XX tie receptor kinase.

XX Tie; receptor, tyrosine kinase, endothelial cell; immunoglobulin; lig;

XX epidermal growth factor; EGF; repeat; blood vessel; atherosclerosis;

XX thromboembolic disease; neoplastic disease; tumour angiogenesis;

XX wound healing; inflammatory disease.

XX Homo sapiens.

XX Key

XX Location/Qualifiers

XX Peptide 1..21

XX /note= "Signal peptide"

XX Protein 22..1138

XX /note= "Mature protein"

XX Modified-site 83..85

XX /note= "N-linked glycosylation site"

XX Modified-site 161..165

FT Domain /note= "N-linked glycosylation site"

FT 214..257

FT /note= "EGF-1"

FT 258..304

FT /note= "EGF-11"

FT 305..346

FT /note= "EGF-111"

FT 503..505

FT /note= "N-linked glycosylation site"

FT 596..598

FT /note= "N-linked glycosylation site"

FT 709..711

FT /note= "N-linked glycosylation site"

FT 761..786

FT /note= "Transmembrane region"

FT 837..1107

FT /note= "Tyrosine kinase domain"

FT 938..951

FT /note= "Kinase insert"

XX W09314124-A.

PN 22-JUL-1993.

XX 08-JAN-1993; 94WO-FI00006.

XX 09-JAN-1992; 92US-0817800.

XX (UYHE-) UNIV HELSINKI HOLDING LTD.

XX Aliitalo K, Armstrong E, Korhonen J, Makela TP;

XX Partanen J;

XX WPI; 1993-243152/30.

XX N-PSDB; AA047064.

XX Nucleic acid encoding tyrosine kinase receptor - for regulating

XX tie activity in diagnosis and treatment of neoplastic diseases

XX involving tumour angiogenesis, wound healing etc.

XX Claim 2; Page 40-44; 73pp; English.

XX The sequence given in AAR39820 represents the full length 'tie' receptor

XX kinase. The sequence given in AAR39821 represents a truncated

XX version of this, wherein residues corresponding to positions 214

XX 257 of the first sequence are absent. Tie is an endothelial cell

XX receptor tyrosine kinase and contains immunoglobulin (Ig-) and

XX epidermal growth factor (EGF-) like repeats. The tie gene is

XX expressed in the endothelial cells of blood vessels. Tie levels may

XX indicate onset/progression of diseases involving endothelial cells

XX and their tie receptors, such as neoplastic diseases involving tumour

XX angiogenesis, wound healing, thromboembolic diseases, atherosclerosis

XX and inflammatory diseases.

XX Sequence 1138 AA;

Query Match 33.2%; Score 865.5; Db 14; Length 1138;

Best Local Similarity 38.2%; Pred No 1 8e-56;

Matches 174; Conservative 71; Mismatches 185; Indels 25; Gaps 8;

QY 14 LLLSGTVEGAMDLILINSPLVSVDAFTSLTCLASQWRHPEPTIGRDPEAL 64

Db 13 tflashvgaavdtltianrltdpqrffltcvsgaaagsgdawqp-pltckddrv 70

QY 65 MNQHQTPIFVTLDVTFEWAQKVMKPEFFASKINAYFEGEVRGEATPTPTMKPQCAFLPATLMTVDRGDNV 124

Db 71 rppppfpr larnqshqtltrqskpsdlqvlccagaaartrtrlychaspah 126

QY 124 FLPATLMTVDRGDNVNSPKKVLKEDEDAVYKNGSFHHSVPREVPDILEVILPHAQPDAGVYSARYIGNLF 180

Db 127 tlpshvrltckadla:lsur:lkexq:dlvksasr:qyltdwheadqrlltq:lpn 184





Genome version 4.5  
Copyright (c) 1993 - 2000 Compaq Ltd.

OM protein - protein search, using sw model

Run on: September 23, 2002, 22:13:40, Search time: 11:02 seconds  
(without alignments)  
724,176 Million cell updates/sec

Title: US-09-733-764-2\_COPY\_1\_472  
Perfect score: 2605  
Sequence: 1 MESA5LVLCVNSLLISCTV VITGRINFAVINTSFFPYS: 472

Scoring table: BLOSUM62  
Gapop 10 0, Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

- Database: Issued Patents.AA.\*
- 1: /cqn2\_6/ptedata/1/1aa/5A.COMB.pep.\*
  - 2: /cqn2\_6/ptedata/1/1aa/5B.COMB.pep.\*
  - 3: /cqn2\_6/ptedata/1/1aa/6A.COMB.pep.\*
  - 4: /cqn2\_6/ptedata/1/1aa/6B.COMB.pep.\*
  - 5: /cqn2\_6/ptedata/1/1aa/PCTUS.COMB.pep.\*
  - 6: /cqn2\_6/ptedata/1/1aa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2605	100.0	1124	1	US-08-323-474-2
2	2605	100.0	1124	5	PCT US93-06093-2
3	2407	92.4	1118	1	US-07-934-393B-2
4	2407	92.4	1118	1	US-08-278-089A-2
5	2407	92.4	1118	2	US-08-838-957A-2
6	2407	92.4	1122	1	US-08-278-089A-6
7	2407	92.4	1122	2	US-08-838-957A-6
8	2352	90.3	1101	2	US-08-469-537A-96
9	865.5	33.2	1138	1	US-08-323-474-8
10	865.5	33.2	1138	2	US-08-469-537A-98
11	865.5	33.2	1138	2	US-08-220-240A-5
12	814	31.2	1135	2	US-08-469-537A-97
13	455	17.5	131	2	US-08-650-598-3
14	305	11.7	294	4	US-09-188-930-332
15	273	11.2	294	4	US-09-188-930-192
16	264	10.4	47	1	US-08-278-089A-14
17	269	10.3	47	2	US-08-838-957A-18
18	242	9.3	44	1	US-08-278-089A-18
19	242	9.3	44	2	US-08-838-957A-17
20	231.5	8.9	685	3	US-08-872-855-2
21	210	8.1	42	1	US-08-278-089A-20
22	203.5	7.8	2471	1	US-08-185-432-16
23	203.5	7.8	2471	1	US-08-083-590A-19
24	203.5	7.8	2471	3	US-08-532-384-19
25	200.5	7.7	1193	2	US-08-400-159-10
26	200.5	7.7	1193	4	US-08-611-729A-10
27	199.5	7.7	1404	2	US-08-400-159-2

RESULT 1  
US-08-323-474-2  
: Sequence 2, Application US/08323474  
: Patent No. 5447860  
: GENERAL INFORMATION:  
: APPLICANT: Ziegler, Steven F.  
: TITLE OF INVENTION: NOVEL TYROSINE KINASE  
: NUMBER OF SEQUENCES: 8  
: CORRESPONDENCE ADDRESS:  
: ADDRESSEE: Immunex Corporation  
: STREET: 51 University Street  
: CITY: Seattle  
: STATE: Washington  
: COUNTRY: US  
: ZIP: 98101  
: COMPUTER READABLE FORM:  
: MEDIUM TYPE: Floppy disk  
: COMPUTER: IBM PC compatible  
: OPERATING SYSTEM: PC-DOS/MS-DOS  
: SOFTWARE: Patent In Release #1.0, Version #1.25  
: CURRENT APPLICATION DATA:  
: APPLICATION NUMBER: US/08/323,474  
: FILING DATE:  
: CLASSIFICATION: 435  
: PRIOR APPLICATION DATA:  
: APPLICATION NUMBER: US 07/905,600  
: FILING DATE: 26-JUN-1992  
: ATTORNEY/AGENT INFORMATION:  
: NAME: Seese, Kathryn A.  
: REGISTRATION NUMBER: 32,172  
: REFERENCE/TITLE NUMBER: 2609  
: TELECOMMUNICATION INFORMATION:  
: TELEPHONE: (206) 587-0430  
: TELEFAX: (206) 233-0644  
: TELEX: 756822  
: INFORMATION FOR SEQ ID NO: 2:  
: SEQUENCE CHARACTERISTICS:  
: LENGTH: 1134 amino acids  
: TYPE: amino acid  
: TOPOLOGY: linear  
: MOLECULE TYPE: protein  
: US 08-323 474-2

ALIGNMENTS

Query Match 100.0%, Score 2605, DB 1, Length 1124;  
Best Local Similarity 100.0%, Pctd. No. 1,3e-217;  
Matches 472, Conservation 0, Mismatches 0, Gaps 0;  
1 MESA5LVLCVNSLLISCTV:VITGRINFAVINTSFFPYS:472  
|||||

Db 1 MDSLASLVGVSLLSSTVFAMDLILINSLPLVSDAETSLTCTASGWRPHEPITIGRD 60  
QY 61 FHALMNOHQDPLEVTOVIVIRWAKKVVWKKRASKINCAVFCGCRVGEAIPRTKMKRQ 120  
Db 61 FHALMNOHQDPLEVTOVIVIRWAKKVVWKKRASKINCAVFCGCRVGEAIPRTKMKRQ 120  
QY 121 QASFLPATLTMVVDKGDVNISFKKVLKEEDAVIYKNGSFTHSVPRHEVPDILEVHLPH 180  
Db 121 QASFLPATLTMVVDKGDVNISFKKVLKEEDAVIYKNGSFTHSVPRHEVPDILEVHLPH 180  
QY 181 AQPDAGVYSARYIGGNFTSAFTPLIVPRFEAUKWGPENHLCTAMNNVCVCHETGEC 240  
Db 181 AQPDAGVYSARYIGGNFTSAFTPLIVPRFEAUKWGPENHLCTAMNNVCVCHETGEC 240  
QY 241 ICPPGFMGRTECKACELHIFKTCNKRCSGQGGCKSVVFCULPDYGGSCALGKGLQCNF 300  
Db 241 ICPPGFMGRTECKACELHIFKTCNKRCSGQGGCKSVVFCULPDYGGSCALGKGLQCNF 300  
QY 301 ACHPGFYGHCKLRCSNNGEMCDRFOGCLGSPQWQGLQCFREGIPRMTPKIVDLPHIE 360  
Db 301 ACHPGFYGHCKLRCSNNGEMCDRFOGCLGSPQWQGLQCFREGIPRMTPKIVDLPHIE 360  
QY 361 VNSGKFNPICKASGWLPTNEEMTLVKPDGTVLHPKDFNHTDHSVAIFTHRIILPPDSG 420  
Db 361 VNSGKFNPICKASGWLPTNEEMTLVKPDGTVLHPKDFNHTDHSVAIFTHRIILPPDSG 420  
QY 421 VVVCSTNTVAGVKEKPFNISVKVLPKPLNAPNVIDTGNFAVINISSEPYEG 472  
Db 421 VVVCSTNTVAGVKEKPFNISVKVLPKPLNAPNVIDTGNFAVINISSEPYEG 472

## RESULT 2

PCT-US93-06093-2  
: Sequence 2, Application PC/US9306093  
: GENERAL INFORMATION:  
: APPLICANT: Ziegler, Steven F.  
: TITLE OF INVENTION: NADEL TYROSINE KINASE  
: NUMBER OF SEQUENCES: 3  
: CORRESPONDENCE ADDRESS:  
: ADDRESSEE: Immunox Corporation  
: STREET: 51 University Street  
: CITY: Seattle  
: STATE: Washington  
: COUNTRY: US  
: ZIP: 98101  
: COMPUTER READABLE FORM:  
: MEDIUM TYPE: Floppy disk  
: COMPUTER: IBM PC compatible  
: OPERATING SYSTEM: PC-DOS/MS-DOS  
: SOFTWARE: Patent In Release #1.0, Version #1.25  
: CURRENT APPLICATION DATA:  
: APPLICATION NUMBER: PCT/US93/06093  
: FILING DATE: 19930625  
: CLASSIFICATION:  
: PRIOR APPLICATION DATA:  
: APPLICATION NUMBER: US 07/905,600  
: FILING DATE: 26-JUN-1992  
: ATTORNEY/AGENT INFORMATION:  
: NAME: Secco, Kathryn A.  
: REGISTRATION NUMBER: 32,172  
: REFERENCE/DOCKET NUMBER: 2609  
: TELECOMMUNICATION INFORMATION:  
: TELEPHONE: (206) 587-0430  
: TELEFAX: (206) 233-0644  
: TELEX: 756422  
: INFORMATION FOR SEQ ID NO: 2:  
: SEQUENCE CHARACTERISTICS:  
: LENGTH: 1124 amino acids  
: TYPE: AMINO ACID  
: TOPOLOGY: linear  
: MOLECULE TYPE: protein  
PCT US93-06093-2

Query Match 100.0%; Score 2605; DB 5; Length 1124;  
Ref Local Similarity 100.0%; Pred No. 1.3e 217;  
Matches 472; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MDSLASLVGVSLLSSTVFAMDLILINSLPLVSDAETSLTCTASGWRPHEPITIGRD 60  
Db 1 MDSLASLVGVSLLSSTVFAMDLILINSLPLVSDAETSLTCTASGWRPHEPITIGRD 60  
QY 61 FHALMNOHQDPLEVTOVIVIRWAKKVVWKKRASKINCAVFCGCRVGEAIPRTKMKRQ 120  
Db 61 FHALMNOHQDPLEVTOVIVIRWAKKVVWKKRASKINCAVFCGCRVGEAIPRTKMKRQ 120  
QY 121 QASFLPATLTMVVDKGDVNISFKKVLKEEDAVIYKNGSFTHSVPRHEVPDILEVHLPH 180  
Db 121 QASFLPATLTMVVDKGDVNISFKKVLKEEDAVIYKNGSFTHSVPRHEVPDILEVHLPH 180  
QY 181 AQPDAGVYSARYIGGNFTSAFTPLIVPRFEAUKWGPENHLCTAMNNVCVCHETGEC 240  
Db 181 AQPDAGVYSARYIGGNFTSAFTPLIVPRFEAUKWGPENHLCTAMNNVCVCHETGEC 240  
QY 241 ICPPGFMGRTECKACELHIFKTCNKRCSGQGGCKSVVFCULPDYGGSCALGKGLQCNF 300  
Db 241 ICPPGFMGRTECKACELHIFKTCNKRCSGQGGCKSVVFCULPDYGGSCALGKGLQCNF 300  
QY 301 ACHPGFYGHCKLRCSNNGEMCDRFOGCLGSPQWQGLQCFREGIPRMTPKIVDLPHIE 360  
Db 301 ACHPGFYGHCKLRCSNNGEMCDRFOGCLGSPQWQGLQCFREGIPRMTPKIVDLPHIE 360  
QY 361 VNSGKFNPICKASGWLPTNEEMTLVKPDGTVLHPKDFNHTDHSVAIFTHRIILPPDSG 420  
Db 361 VNSGKFNPICKASGWLPTNEEMTLVKPDGTVLHPKDFNHTDHSVAIFTHRIILPPDSG 420  
QY 421 VVVCSTNTVAGVKEKPFNISVKVLPKPLNAPNVIDTGNFAVINISSEPYEG 472  
Db 421 VVVCSTNTVAGVKEKPFNISVKVLPKPLNAPNVIDTGNFAVINISSEPYEG 472

## RESULT 3

US-07-934-393H-2  
: Sequence 2, Application US/07934393H  
: Patent No. 5466596  
: GENERAL INFORMATION:  
: APPLICANT: BREITMAN, MARTIN L.  
: APPLICANT: DUMONT, DANIEL  
: APPLICANT: GRADWOHL, GERARD G.  
: TITLE OF INVENTION: ISSUE SPECIFIC TRANSCRIPTIONAL  
: TITLE OF INVENTION: REGULATORY ELEMENT  
: NUMBER OF SEQUENCES: 5  
: CORRESPONDENCE ADDRESS:  
: ADDRESSEE: HERESKIN & PARK  
: STREET: 40 King Street West  
: CITY: Toronto  
: STATE: Ontario  
: COUNTRY: Canada  
: ZIP: M5H 3Y2  
: COMPUTER READABLE FORM:  
: MEDIUM TYPE: Floppy disk  
: COMPUTER: IBM PC compatible  
: OPERATING SYSTEM: PC-DOS/MS-DOS  
: SOFTWARE: Patent In Release #1.0, Version #1.25  
: CURRENT APPLICATION DATA:  
: APPLICATION NUMBER: US/07934,393H  
: FILING DATE: 25 AUG 1992  
: CLASSIFICATION: 435  
: ATTORNEY/AGENT INFORMATION:  
: NAME: Kurdydk, Linda M.  
: REGISTRATION NUMBER: 34,971  
: REFERENCE/DOCKET NUMBER: 3153-64  
: TELECOMMUNICATION INFORMATION:  
: TELEPHONE: (416) 354-7311  
: TELEFAX: (416) 361-1398  
: INFORMATION FOR SEQ ID NO: 2:

```

: SEQUENCE CHARACTERISTICS:
: LENGTH: 1118 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
US-07-934-393B-2

```

```

Query Match: 92.4%; Score 2407; DH 1; length 1118;
Best local Similarity 91.9%; Pred. No. 2e-200;
Matches 434; Conservative 17; Mismatches 21; Indels 0; Gaps 0;

```

```

QY 1 MDSLASLVLCGVSLLSSTVSGAMDLILINSUPLVSDAETSLTCTIASWHPHEPITIGRD 60
DB 1 MDSLAGLVLCGVSLLSYGVVEGAMDLILINSUPLVSDAETSLTCTIASWHPHEPITIGRD 60
QY 61 FEALMNOHQDPLEVTCQVTFEWAQKVVWKKRASKINAYPCECHVPGCAIPIPTMKMPQ 120
DB 61 FEALMNOHQDPLEVTCQVTFEWAQKVVWKKRASKINAYPCECHVPGCAIPIPTMKMPQ 120
QY 121 QASFLPATILMTVDKGNVNSFPKKVLIKEDAVIYKNGSFTHSVPRHEVPDILKVLHLP 180
DB 121 QASFLPATILMTVDKGNVNSFPKKVLIKEDAVIYKNGSFTHSVPRHEVPDILKVLHLP 180
QY 181 AQPDAGVVSARYIGNLFTSAFTPLIVPPEALQKWGPENHLCTACMNNGVCHEDTGE 240
DB 181 AQPDAGVVSARYIGNLFTSAFTPLIVPPEALQKWGPENHLCTACMNNGVCHEDTGE 240
QY 241 ICPPGFMGRTCERACELHTFGRTCKEPCSGQGGCKSYVFCGLDPYGCSCATGKGLQNE 300
DB 241 ICPPGFMGRTCERACELHTFGRTCKEPCSGQGGCKSYVFCGLDPYGCSCATGKGLQNE 300
QY 301 ACHPGYGPDPCKLRCSCNNGCMCDRPGGCGSPGQGGLOCHREGIPRMTPKIVDLPDHIE 360
DB 301 ACHPGYGPDPCKLRCSCNNGCMCDRPGGCGSPGQGGLOCHREGIPRMTPKIVDLPDHIE 360
QY 361 VNSCKFNPKCKASGWPPLPTNEEMTLVKPDGTVLHPKDNHIDHFSVAITFTHRIILPPDSG 420
DB 361 VNSCKFNPKCKASGWPPLPTNEEMTLVKPDGTVLHPKDNHIDHFSVAITFTHRIILPPDSG 420
QY 421 VWCVSVNTVAGMVEKPFNISVKVLKPLNAPNVITDGHFAVINISSEPYFG 472
DB 421 VWCVSVNTVAGMVEKPFNISVKVLKPLNAPNVITDGHFAVINISSEPYFG 472

```

## RESULT 4

```

: Sequence 2, Application US/08278089A
: Patent No. 5681714
: GENERAL INFORMATION:
: APPLICANT: Breitman, Martin L.
: APPLICANT: Rossant, Janet
: APPLICANT: Dumont, Daniel J.
: APPLICANT: Yamauchi, Terry P.
: TITLE OF INVENTION: No. 5681714e1 Receptor Tyrosine Kinase
: NUMBER OF SEQUENCES: 33
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Hereskin & Parr
: STREET: 40 King Street West
: CITY: Toronto
: STATE: Ontario
: COUNTRY: Canada
: ZIP: M5H 3Y2
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent In Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/278,089A
: FILING DATE: 20-JUL-1994
: CLASSIFICATION: 530
: ATTORNEY/AGENT INFORMATION:

```

```

: NAME: Kurdzyk, Linda M.
: REGISTRATION NUMBER: 34,971
: REFERENCE/DOCKET NUMBER: 3153-111
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (416) 364-7311
: TELEFAX: (416) 361-1398
: INFORMATION FOR SEQ ID NO: 2:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1118 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
US-08-278-089A-2

```

```

Query Match: 92.4%; Score 2407; DH 1; length 1118;
Best local Similarity 91.9%; Pred. No. 2e-200;
Matches 434; Conservative 17; Mismatches 21; Indels 0; Gaps 0;

```

```

QY 1 MDSLASLVLCGVSLLSSTVSGAMDLILINSUPLVSDAETSLTCTIASWHPHEPITIGRD 60
DB 1 MDSLAGLVLCGVSLLSYGVVEGAMDLILINSUPLVSDAETSLTCTIASWHPHEPITIGRD 60
QY 61 FEALMNOHQDPLEVTCQVTFEWAQKVVWKKRASKINAYPCECHVPGCAIPIPTMKMPQ 120
DB 61 FEALMNOHQDPLEVTCQVTFEWAQKVVWKKRASKINAYPCECHVPGCAIPIPTMKMPQ 120
QY 121 QASFLPATILMTVDKGNVNSFPKKVLIKEDAVIYKNGSFTHSVPRHEVPDILKVLHLP 180
DB 121 QASFLPATILMTVDKGNVNSFPKKVLIKEDAVIYKNGSFTHSVPRHEVPDILKVLHLP 180
QY 181 AQPDAGVVSARYIGNLFTSAFTPLIVPPEALQKWGPENHLCTACMNNGVCHEDTGE 240
DB 181 AQPDAGVVSARYIGNLFTSAFTPLIVPPEALQKWGPENHLCTACMNNGVCHEDTGE 240
QY 241 ICPPGFMGRTCERACELHTFGRTCKEPCSGQGGCKSYVFCGLDPYGCSCATGKGLQNE 300
DB 241 ICPPGFMGRTCERACELHTFGRTCKEPCSGQGGCKSYVFCGLDPYGCSCATGKGLQNE 300
QY 301 ACHPGYGPDPCKLRCSCNNGCMCDRPGGCGSPGQGGLOCHREGIPRMTPKIVDLPDHIE 360
DB 301 ACHPGYGPDPCKLRCSCNNGCMCDRPGGCGSPGQGGLOCHREGIPRMTPKIVDLPDHIE 360
QY 361 VNSCKFNPKCKASGWPPLPTNEEMTLVKPDGTVLHPKDNHIDHFSVAITFTHRIILPPDSG 420
DB 361 VNSCKFNPKCKASGWPPLPTNEEMTLVKPDGTVLHPKDNHIDHFSVAITFTHRIILPPDSG 420
QY 421 VWCVSVNTVAGMVEKPFNISVKVLKPLNAPNVITDGHFAVINISSEPYFG 472
DB 421 VWCVSVNTVAGMVEKPFNISVKVLKPLNAPNVITDGHFAVINISSEPYFG 472

```

## RESULT 5

```

: Sequence 2, Application US/08838957A
: Patent No. 5948187
: GENERAL INFORMATION:
: APPLICANT: Breitman, Martin L.
: APPLICANT: Rossant, Janet
: APPLICANT: Dumont, Daniel J.
: APPLICANT: Yamauchi, Terry P.
: TITLE OF INVENTION: No. 5948187e1 Receptor Tyrosine Kinase
: NUMBER OF SEQUENCES: 32
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Hereskin & Parr
: STREET: 40 King Street West
: CITY: Toronto
: STATE: Ontario
: COUNTRY: Canada
: ZIP: M5H 3Y2
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible

```

? OPERATING SYSTEM: PC-DOS/MS-DOS  
 ? SOFTWARE: Patent In Release #1 0, Version #1.30  
 ? CURRENT APPLICATION DATA:  
 ? APPLICATION NUMBER: US/09/638,957A  
 ? FILING DATE: 23-APR-1997  
 ? CLASSIFICATION: 435  
 ? ATTORNEY/AGENT INFORMATION:  
 ? NAME: Kurdydyk, Linda M.  
 ? REGISTRATION NUMBER: 34,971  
 ? REFERENCE/DOCKET NUMBER: 4153-212  
 ? TELECOMMUNICATION INFORMATION:  
 ? TELEPHONE: (416) 364-7311  
 ? TELEFAX: (416) 361-1398  
 ? INFORMATION FOR SEQ ID NO: 2:  
 ? SEQUENCE CHARACTERISTICS:  
 ? LENGTH: 1118 amino acids  
 ? TYPE: amino acid  
 ? TOPOLOGY: linear  
 ? MOLECULE TYPE: protein  
 ? US 08-838-957A-2

Query Match 92.4%; Score 2407; DB 2; length 1118,  
 Best Local Similarity 91.9%; Fred. No. 2e-200;  
 Matches 434; Conservative 17; Mismatches 21; Indels 0; Gaps 0;  
 QY 1 MSLASLVLCGVSHLSCTVGCAMDLLINSLPLVSDAFTSLICIASCWPHPEPTIGRD 60  
 DB 1 MSLASLVLCGVSHLSCTVGCAMDLLINSLPLVSDAFTSLICIASCWPHPEPTIGRD 60  
 QY 61 FEALNNHQDPLEVTDVTPWAKKVVWKKKASKINAYFCGVPVSEAIRITPKMPQ 120  
 DB 61 FEALNNHQDPLEVTDVTPWAKKVVWKKKASKINAYFCGVPVSEAIRITPKMPQ 120  
 QY 121 QASFLPATLTMTVDKGNVNI SFKKVLIKEDAVIYKNGSFTHSVPRHVPDILEVILPH 180  
 DB 121 QASFLPATLTMTVDKGNVNI SFKKVLIKEDAVIYKNGSFTHSVPRHVPDILEVILPH 180  
 QY 181 AQPDAGVYSARYIGGNLFTSAFTPLIVRRCEAKQWGPCNHLCTACMNNVCVCHETGEC 240  
 DB 181 AQPDAGVYSARYIGGNLFTSAFTPLIVRRCEAKQWGPCNHLCTACMNNVCVCHETGEC 240  
 QY 241 ICPPGFMGRTEKACELHITFRTCKERGSGDEGKSVVECLPQYGSQATKWKGLQNE 300  
 DB 241 ICPPGFMGRTEKACELHITFRTCKERGSGDEGKSVVECLPQYGSQATKWKGLQNE 300  
 QY 301 ACPPGFMGRTEKACELHITFRTCKERGSGDEGKSVVECLPQYGSQATKWKGLQNE 360  
 DB 301 ACPPGFMGRTEKACELHITFRTCKERGSGDEGKSVVECLPQYGSQATKWKGLQNE 360  
 QY 361 VNSGKFNPICKASQWPLTNEMLVKHGDGIVLHPKDFNHDHFSVAIFTTHRLIPDSC 420  
 DB 361 VNSGKFNPICKASQWPLTNEMLVKHGDGIVLHPKDFNHDHFSVAIFTTHRLIPDSC 420  
 QY 421 VWCVSVNTVACWVKRPNISVKVLPEPLHAPNVDIGHNFAINISSEPEFG 472  
 DB 421 VWCVSVNTVACWVKRPNISVKVLPEPLHAPNVDIGHNFAINISSEPEFG 472

RESULT 6  
 US 08-838-957A-6  
 ? Sequence 6, Application US/09/638,957A  
 ? Patent No. 5681714  
 ? GENERAL INFORMATION:  
 ? APPLICANT: Breitman, Martin L.  
 ? APPLICANT: Rossant, Janet  
 ? APPLICANT: Dumont, Daniel J.  
 ? APPLICANT: Yamashita, Terry P.  
 ? TITLE OF INVENTION: No. 5681714c1 Receptor Tyrosine Kinase  
 ? NUMBER OF SEQUENCES: 33  
 ? CORRESPONDENCE ADDRESS:  
 ? ADDRESSEE: Breitman, Martin L.  
 ? STREET: 40 King Street West

? CITY: Toronto  
 ? STATE: Ontario  
 ? COUNTRY: Canada  
 ? ZIP: M5H 1Y2  
 ? COMPUTER READABLE FORM:  
 ? MEDIUM TYPE: Floppy disk  
 ? COMPUTER: IBM PC compatible  
 ? OPERATING SYSTEM: PC-DOS/MS-DOS  
 ? SOFTWARE: Patent In Release #1 0, Version #1 40  
 ? CURRENT APPLICATION DATA:  
 ? APPLICATION NUMBER: US/08/278,089A  
 ? FILING DATE: 20-JUL-1994  
 ? CLASSIFICATION: 530  
 ? ATTORNEY/AGENT INFORMATION:  
 ? NAME: Kurdydyk, Linda M.  
 ? REGISTRATION NUMBER: 34,971  
 ? REFERENCE/DOCKET NUMBER: 3153-111  
 ? TELECOMMUNICATION INFORMATION:  
 ? TELEPHONE: (416) 364-7311  
 ? TELEFAX: (416) 361-1398  
 ? INFORMATION FOR SEQ ID NO: 6:  
 ? SEQUENCE CHARACTERISTICS:  
 ? LENGTH: 1122 amino acids  
 ? TYPE: amino acid  
 ? TOPOLOGY: linear  
 ? MOLECULE TYPE: protein  
 ? US-08-278-089A-6

Query Match 92.4%; Score 2407; DB 1; Length 1122;  
 Best Local Similarity 91.9%; Fred. No. 2e-200;  
 Matches 434; Conservative 17; Mismatches 21; Indels 0; Gaps 0;  
 QY 1 MSLASLVLCGVSHLSCTVGCAMDLLINSLPLVSDAFTSLICIASCWPHPEPTIGRD 60  
 DB 1 MSLASLVLCGVSHLSCTVGCAMDLLINSLPLVSDAFTSLICIASCWPHPEPTIGRD 60  
 QY 61 FEALNNHQDPLEVTDVTPWAKKVVWKKKASKINAYFCGVPVSEAIRITPKMPQ 120  
 DB 61 FEALNNHQDPLEVTDVTPWAKKVVWKKKASKINAYFCGVPVSEAIRITPKMPQ 120  
 QY 121 QASFLPATLTMTVDKGNVNI SFKKVLIKEDAVIYKNGSFTHSVPRHVPDILEVILPH 180  
 DB 121 QASFLPATLTMTVDKGNVNI SFKKVLIKEDAVIYKNGSFTHSVPRHVPDILEVILPH 180  
 QY 181 AQPDAGVYSARYIGGNLFTSAFTPLIVRRCEAKQWGPCNHLCTACMNNVCVCHETGEC 240  
 DB 181 AQPDAGVYSARYIGGNLFTSAFTPLIVRRCEAKQWGPCNHLCTACMNNVCVCHETGEC 240  
 QY 241 ICPPGFMGRTEKACELHITFRTCKERGSGDEGKSVVECLPQYGSQATKWKGLQNE 300  
 DB 241 ICPPGFMGRTEKACELHITFRTCKERGSGDEGKSVVECLPQYGSQATKWKGLQNE 300  
 QY 301 ACPPGFMGRTEKACELHITFRTCKERGSGDEGKSVVECLPQYGSQATKWKGLQNE 360  
 DB 301 ACPPGFMGRTEKACELHITFRTCKERGSGDEGKSVVECLPQYGSQATKWKGLQNE 360  
 QY 361 VNSGKFNPICKASQWPLTNEMLVKHGDGIVLHPKDFNHDHFSVAIFTTHRLIPDSC 420  
 DB 361 VNSGKFNPICKASQWPLTNEMLVKHGDGIVLHPKDFNHDHFSVAIFTTHRLIPDSC 420  
 QY 421 VWCVSVNTVACWVKRPNISVKVLPEPLHAPNVDIGHNFAINISSEPEFG 472  
 DB 421 VWCVSVNTVACWVKRPNISVKVLPEPLHAPNVDIGHNFAINISSEPEFG 472

RESULT 7  
 US-08-838-957A-6  
 ? Sequence 6, Application US/08/278,089A  
 ? Patent No. 5998187  
 ? GENERAL INFORMATION:  
 ? APPLICANT: Breitman, Martin L.  
 ? APPLICANT: Rossant, Janet



APPLICANT: Dumont, Daniel J.  
 APPLICANT: Yamauchi, Terry P.  
 TITLE OF INVENTION: No. 598187e1 Receptor Tyrosine Kinase  
 NUMBER OF SEQUENCES: 32  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Bereskin & Parr  
 CITY: Toronto  
 STATE: Ontario  
 COUNTRY: Canada  
 ZIP: M5H 3Y2  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/848,957A  
 FILING DATE: 23 APR 1997  
 CLASSIFICATION: 435  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Kutydyk, Linda M.  
 REGISTRATION NUMBER: 34,971  
 REFERENCE/DOCKET NUMBER: 3153-212  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (416) 364-7311  
 TELEFAX: (416) 361-1398  
 INFORMATION FOR SEQ ID NO: 6:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 1122 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 US-08-48-957A-6

Query Match 92.4%, Score 2407, DB 2; Length 1122.  
 Best Local Similarity 91.9%; Pred. No. 2e-200;  
 Matches 434; Conservative 17; Mismatches 21; Indels 0; Gaps 0;

QY 1 MDSIASVLGCVSLLSGTVHGAMDLLINSLPLVSDAETSLTCTASGWRPHEPITIGRD 60  
 DB 1 MDSLAGLVLCGSLLLGVVEGAMDLLINSLPLVSDAETSLTCTASGWRPHEPITIGRD 60  
 QY 61 FEALMNQHDPLVTVQVTRWAKKVVWKKREKASKINGAYFCEGRVGRGAIRITMKMPQ 120  
 DB 61 FEALMNQHDPLVTVQVTRWAKKVVWKKREKASKINGAYFCEGRVGRGAIRITMKMPQ 120  
 QY 121 QASLPATLITMDKGDVNISFKKVLKEEDAVIYKNGSFTHSVPRHEVPDILEVHLPH 180  
 DB 121 QASLPATLITMDKGDVNINSEKVLKEEDAVIYKNGSFTHSVPRHEVPDILEVHLPH 180  
 QY 181 ACPGACVYSARYIDENIFSAITPIIVVWTPAKKWQPEINHTCTACMNKQVCHETDCE 240  
 DB 181 ACPGACVYSARYIGNLTSAITPLIVRPPEAKKWQPEINHTCTACMNKQVCHETDCE 240  
 QY 241 ICPFGMGRTCEKACELHRTGCKRCSCGPGCKSVYFCLDPYGCSCAIQWKGLQCNF 300  
 DB 241 ICPFGMGRTCEKACELHRTGCKRCSCGPGCKSVYFCLDPYGCSCAIQWKGLQCNF 300  
 QY 301 AITHEGEGSPKLRQSSNNSEMTRFQWGLSPKWAQGLQERESIPPMTPKIVGLPDHLE 360  
 DB 301 ACPGSGYCPDCKLRCHCINERICDRPGGCLCSQGWGLQCKKRGKPRMTIPIQEDLPDHLF 360  
 QY 361 VNSKFNPICKASQWPLPIPHMTLVKPGDTVLHPKDFNTHDFSAITTHRLPPDSG 420  
 DB 361 VNSKFNPICKASQWPLPIPHMTLVKPGDTVLHPKDFNTHDFSAITTHRLPPDSG 420  
 QY 421 VVGSVNTVAGMVEKPNISVKVLPIKPLINAPNVLDTGHNFAVINISSEPYFG 472  
 DB 421 VVGSVNTVAGMVEKPNISVKVLPIKPLINAPNVLDTGHNFAVINISSEPYFG 472

RESULT 8  
 US-08-469-537A-96  
 Sequence 96, Application US/08469537A  
 Patent No. 5843749  
 GENERAL INFORMATION:  
 APPLICANT: Maisongierre, et al.  
 TITLE OF INVENTION: EHK AND FOR TYROSINE  
 TITLE OF INVENTION: KINASES  
 NUMBER OF SEQUENCES: 107  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Regeneron Pharmaceuticals, Inc.  
 STREET: 777 Old Saw Mill River Road  
 CITY: Tarrytown  
 STATE: NY  
 COUNTRY: U.S.A.  
 ZIP: 10591  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Diskette  
 COMPUTER: IBM compatible  
 OPERATING SYSTEM: DOS  
 SOFTWARE: PstSPQ Version 2.0  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/469,537A  
 FILING DATE: 06-JUN-1995  
 CLASSIFICATION: 435  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: USSN 08/406,247  
 FILING DATE: 17-MAR-1995  
 APPLICATION NUMBER: USSN 08/144,992  
 FILING DATE: 28-OCT-1993  
 APPLICATION NUMBER: USSN 07/736,559  
 FILING DATE: 26-JUL-1991  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Kemper, Ph.D., Gail M  
 REGISTRATION NUMBER: 32,143  
 REFERENCE/DOCKET NUMBER: REG 070C  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 914-345-7400  
 TELEFAX: 914-345-7721  
 TELEX:  
 INFORMATION FOR SEQ ID NO: 96:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 1101 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: unknown  
 TOPOLOGY: unknown  
 MOLECULE TYPE: protein  
 US-08-469-537A-96

Query Match 90.3%, Score 2352; DB 2; Length 1101;  
 Best Local Similarity 93.4%; Pred. No. 1.2e-195;  
 Matches 423; Conservative 12; Mismatches 18; Indels 0; Gaps 0;

QY 20 VEGAMDILLINSLPLVSDAETSLTCTASGWRPHEPITIGRDFAALMNQHDPLEVTDVY 79  
 DB 1 VEGAMDILLINSLPLVSDAETSLTCTASGWRPHEPITIGRDFAALMNQHDPLEVTDVY 60  
 QY 80 REWAKKVVWKKREKASKINGAYFCEGRVGRGAIRITMKKROQASFLPATLITMTVHKDNY 139  
 DB 61 REWAKKVVWKKREKASKINAYFCEGRVGRGAIRITMKKROQASFLPATLITMTVHKDNY 120  
 QY 140 NISPKKVLKEEDAVIYKNGSFTHSVPRHEVPDILEVHLPHAQGPQVAGYSARYIGCNLF 199  
 DB 121 NISPKKVLKEEDAVIYKNGSFTHSVPRHEVPDILEVHLPHAQGPQVAGYSARYIGCNLF 180  
 QY 200 TSAETPLIVRPPEAKKWQPEINHTCTACMNKQVCHETDCEICPPFGMGRTCEKACELHT 259  
 DB 181 TSAETPLIVRPPEAKKWQPEINHTCTACMNKQVCHETDCEICPPFGMGRTCEKACELHT 240  
 QY 260 FGRICKRCSCGPGCKSVYFCLDPYGCSCAIQWKGLQCNFAPHPGFGYGHMKIPKPSNN 319  
 DB 241 FGRICKRCSCGPGCKSVYFCLDPYGCSCAIQWKGLQCNFAPHPGFGYGHMKIPKPSNN 300





```

? COMPUTER: IBM Compatible
? OPERATING SYSTEM: DOS
? SOFTWARE: FASTSEQ Version 2.0
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/38469,537A
? FILING DATE: 06-JUN-1995
? CLASSIFICATION: 435
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: USSN 08/406,247
? FILING DATE: 17-MAR-1995
? APPLICATION NUMBER: USSN 08/144,992
? FILING DATE: 28-OCT-1993
? APPLICATION NUMBER: USSN 07/736,559
? FILING DATE: 26-JUL-1991
? ATTORNEY/AGENT INFORMATION:
? NAME: Kempler, Ph.D., Gail M
? REGISTRATION NUMBER: 32,143
? REFERENCE/DOCKET NUMBER: REG 070C
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: 914-345-7400
? TELEFAX: 914-345-7721
?
? INFORMATION FOR SEQ ID NO: 57:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 1145 amino acids
? TYPE: amino acid
? STRANDEDNESS: unknown
? TOPOLOGY: unknown
? MOLECULE TYPE: protein
? US 08 469-537A-97

```

```

Query Match 31.2%, Score 814, DB 2, Length 1135;
Best Local Similarity 37.2%, Pred. No. 3,4e-62;
Matches 169; Conservative 70; Mismatches 191; Indels 24; Gaps 10;

QY 14 LLLSGIVGCMDELLLSPLVLSIAFTSLICIASCWPRPHTTIGRDFEA---LMOHOD 70
DB 13 LFLSHVAGSDLLTLLNLTETEPQRFLLICVSG-----EAGGRSSDVMARLLLEKDD 66

QY 71 PLEVT-----QDVTRWAKKVVWKP-FKASKINGAYFEISVPVGEAIPRTMKMQQAS 123
DB 67 RIVRFPFPGQPLHTRNGSHQVTLRGSKSDLVGVFSCVGGACTRPTVIVVHNSGAAH 126

QY 124 FLPATLTWVKGNNVNISEKVLKEEDAVIYKNSFIISVPRHEVPDI-LEVHLPAAQ 182
DB 127 LEPDQVTTIVKNGDVAVLARVHKQVTDVIWKNNGSYFHTLDMHEAHDAAGELQNVQ 186

QY 183 PDQAGVYSARYIGCNLTFSATFRLIVKCLAAKQWCPKCNHICLACMNGVCHEDTGRCIC 242
DB 187 PPSAIAVATYLEASPF-ELEFLIVRGTENGSPWPGVCKGTPWGLUSGVCHHDEGVC 245

QY 243 PPRPMWRP-ERACELHTRFTCKERCSDGCKSYVPCLPDPYGCSC/ATGKGLQCNAC 302
DB 246 PPRFTFTFTFQACPRGPRGQSCQPPQCTAGCGRIPTCLPDPYGCSCGCGWPGSCQCEAC 305

QY 303 BRGEVGPD-KLPQSN-NN-DEM-CRPGVY-LQSP-WALCL-TEVEGLIPRATPKIVLPQHLEVN 362
DB 306 ALCHEGAD-RLQVQ-UNQVY-TRFSR-PTWPSGWHIVHCEKSD---PIPQLNVAILELFQ 362

QY 363 SCKPNPI-CKAGWPLPIINERMLIKVDCG-VLHPKDFNHIDFESVAIFTHRIIDPQSC 420
DB 363 LKTMPRINCAANGPFPVPGSMELPKPDCTMIESTRAIVFPDP-TTAEFVPPPLTIADSG 421

QY 421 VVWCVSVNTVAGWKEKPNISVKVLPKLNAPNVI 454
DB 422 FWECKVSTSGGQDSRRKRVNVVIVPLTAHRL 455

```

```

RESULT 13
US-08-650-598-3
? Sequence 1, Application US/08650598
? Patent No. 5877020

```

```

? GENERAL INFORMATION:
? APPLICANT: Alitalo, Karl
? TITLE OF INVENTION: Promoter of the Receptor Tyrosine Kinase, TIE
? NUMBER OF SEQUENCES: 11
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: Marshall, O'Toole, Gotstein, Murray & Horuh
? STREET: 6300 Sears Tower, 233 South Wacker Drive
? CITY: Chicago
? STATE: Illinois
? COUNTRY: United States of America
? ZIP: 60606-6402
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? COMPUTER: IBM PC Compatible
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: PatentIn Release #1.0, Version #1.25
? CURRENT APPLICATION DATA: US/08/650,598
? FILING DATE:
? CLASSIFICATION:
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: 08/110,717
? FILING DATE: 22-SEP-1994
? ATTORNEY/AGENT INFORMATION:
? NAME: Gass, David A.
? REGISTRATION NUMBER: 38,153
? REFERENCE/DOCKET NUMBER: 28113/33245
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: 312/474-6300
? TELEFAX: 312/474-0448
? TELEEX: 25-3856
? INFORMATION FOR SEQ ID NO: 3:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 131 amino acids
? TYPE: amino acid
? STRANDEDNESS: single
? TOPOLOGY: linear
? MOLECULE TYPE: protein
? US-08-650-598-3

```

```

Query Match 17.5%, Score 455, DB 2, Length 131;
Best Local Similarity 53.1%, Pred. No. 2,9e-42;
Matches 69; Conservative 21; Mismatches 38; Indels 2; Gaps 2;

QY 226 -NHLDVTA-MNN-VW-HEIDREY-PPFPHM-PPFKAVALHITFGTCKERSGEGTKSYV 279
DB 2 CVKEDPQTHHGV-HHMG-NVCPQG-IGTPQGLACDFGFGLSQCOQCTGACGCLIF 60

QY 280 -LPIPPYGS-CAIGWGL-ENH-ACHPGYGGPKKPCNNNGEMEDPFGGCLSSPGWGLQ 339
DB 61 -LPIPPYGS-STSS-WPSS-GLQEA-APGHP-GLQ-PLQV-GLQ-NHSD-TPPSS-V-PS-MSBGM 120

QY 340 GR-EGIPRM 348
DB 121 CKSDRIPIQI 130

RESULT 14
US-09-188-930-332
? Sequence 332, Application US/09188930A
? Patent No. 6150502
? GENERAL INFORMATION:
? APPLICANT: Watson, James D.
? APPLICANT: Strachan, Lorna
? APPLICANT: Sleeman, Matthew
? APPLICANT: Gurust, Rene
? APPLICANT: Murison, James Greg
? TITLE OF INVENTION: Compositions Isolated from Skin Cells
? TITLE OF INVENTION: and Methods For Their Use
? FILE REFERENCE: 11000.1011c1
? CURRENT APPLICATION NUMBER: US/09/188,940A
? CURRENT FILING DATE: 1998-11-09

```





GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: September 23, 2002, 23:13:04 ; Search time 20.96 Seconds  
(without alignments)  
2163.844 Million cell updates/sec

Title: us-09-733-764-2\_copy\_1\_472  
Perfect score: 2605

Sequence: 1 MDSIASLVGVSLLSSTV

VILP:HNFAVINISSEPPV: 472

Scoring table: RGSIM62  
Gapop 10.0, Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: pir1:\*

2: pir2:\*

3: pir3:\*

4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution

# SUMMARIES

Result No	Score	Query Match	Length	PF	ID	Description
1	2605	100	9	1124	1	protein-tyrosine k
2	2435	93	5	1126	1	protein-tyrosine k
3	2407	92	4	1122	2	protein-tyrosine k
4	2407	92	4	1123	1	protein-tyrosine k
5	2337	89	7	1125	1	protein-tyrosine k
6	873	5	33	1146	1	protein-tyrosine k
7	865	5	33	1138	1	protein-tyrosine k
8	820	5	31	1134	1	protein-tyrosine k
9	321	12	3	1620	2	protein-tyrosine k
10	317	5	12	1574	2	hypothetical prote
11	280	10	7	1111	2	MEME protein - ra
12	231	5	8	688	2	hypothetical prote
13	221	8	5	3672	2	Delta-4 protein
14	221	8	5	3704	2	hypothetical prote
15	212	8	1	3712	2	hypothetical prote
16	205	5	7	1204	2	laminin alpha 1 ch
17	205	5	7	1964	2	Metch R protein -
18	200	5	7	2471	2	notch4 - mouse
19	199	5	7	686	2	cell-tate determi
20	199	5	7	1408	2	beta-4 protein
21	198	5	7	474	2	gene seriate prote
22	198	5	7	832	2	wnt inhibitory fac
23	198	5	7	833	2	neurogenic protein
24	198	5	7	880	2	gene Delta protein
25	194	5	7	2825	2	neurogenic repetit
26	194	5	7	114271	2	beta-4 protein, str
27	193	5	7	1722	2	protein P11C7.4 li
28	193	5	7	1687	2	EGF repeat transm
29	192	5	7	2555	2	notch protein homo
30	192	5	7	722	2	DELTA-like 1 mem

30 192.5 7.4 1429 2 S06434 homeotic protein 1  
31 191.5 7.4 2524 2 A35844 Xotch protein - Af  
32 191.5 7.4 3635 2 T10053 laminin alpha 5 ch  
33 189.5 7.3 728 2 S0719 C-beta1 - chicke  
34 189.5 7.3 2180 2 S29764 hypothetical prote  
35 189 7.3 473 2 A56195 adhesive plaque pr  
36 189 7.3 1625 2 T42526 secreted leucine-r  
37 188 7.2 1220 2 A56136 jagged protein pre  
38 187.5 7.2 378 2 S59180 wnt inhibitory fac  
39 187.5 7.2 1531 2 T42218 slit-1 protein hom  
40 187.5 7.2 2318 2 S45306 notch 3 protein -  
41 186.5 7.2 1372 2 T25933 hypothetical prote  
42 185.5 7.1 2321 2 S78549 notch3 protein - h  
43 185.5 7.1 2703 2 A54420 notch protein - ir  
44 184.5 7.1 861 2 A48825 Notch homolog Motc  
45 184.5 7.1 1810 1 A32230 tenascin precursor

## ALIGNMENTS

RESULT 1  
158388  
protein-tyrosine kinase (EC 2.7.1.112), receptor type tek precursor - human  
C:Species: Homo sapiens (man)  
C:Date: 02-Jul-1998 #sequence\_revision 02 Jul 1998 #re-charge 16 Jul 1999  
C:Accession: 158388  
R:Ziegler, S.F.; Bird, T.A.; Scherfing, J.A.; Scholley, K.A.; Baum, P.R.  
Oncogene 8, 663-670, 1993  
A:Title: Molecular cloning and characterization of a novel receptor protein tyrosine  
A:Reference number: 158388; MUID:93173509  
A:Accession: 158388  
A:Status: preliminary; translated from GH/EMBL/DDRHJ  
A:Molecule type: mRNA  
A:Residues: 1-1124 <RES>  
A:Cross-references: GB:L19139; NID:q292833; FID:AAA61139.1; FID:q292834  
C:Genetics:  
A:Gene: GDR-TEK  
A:Cross-references: GDR:344185; OMIM:600221  
A:Map position: 9p21-9p21  
C:Function:  
A:Description: catalyzes the phosphorylation of a peptidyl tyrosine residue by ATP  
C:Superfamily: protein-tyrosine kinase, receptor type tek; EGF homolog; fibronectin  
C:Keywords: ATP, autophosphorylation, duplication, glycoprotein, phosphoprotein; phos  
F:1-22/Domain: signal sequence #status predicted <SIG>  
F:23-1124/Product: protein-tyrosine kinase, predicted <SIG>  
F:37-194/Domain: immunoglobulin homology <IM>  
F:135-137/Region: cell attachment (R-G-D) motif  
F:21: 21/Domain: EGF homology <EGD>  
F:255-298/Domain: EGF homology <EGD>  
F:302-340/Domain: EGF homology <EGD>  
F:364-426/Domain: immunoglobulin homology <IM>  
F:447-527/Domain: fibronectin type III repeat homology <FN3>  
F:542-635/Domain: fibronectin type III repeat homology <FN3>  
F:638-720/Domain: fibronectin type III repeat homology <FN3>  
F:752-775/Domain: transmembrane #status predicted <IM>  
F:822-1093/Domain: protein kinase homology <PK>  
F:830-838/Domain: protein kinase ATP binding motif  
F:140-158,399,438,464,560,576,649,691/Binding site: catalyze (Asn) (covalent) #st  
F:865,872,964/Active site: Lys, Glu, Asp #status predicted

Query Match: 100.0% Score: 2605, Len: 1, Length: 1124;

Best Local Similarity: 100.0%, P: 0.0, No. Hits: 0, Gaps: 0;  
Matches: 472, Conservative: 0, Mismatches: 0, Gaps: 0;

QY 1 MDSIASLVGVSLLSSTV:AMDLILINSPLVSDAETSLTIIASQWPPHEPTIGRD 60

DB 1 MDSIASLVGVSLLSSTV:AMDLILINSPLVSDAETSLTIIASQWPPHEPTIGRD 60

QY 61 FTATVNHCHGPPLEVTQVTPFWAKVWVWPPKASHINZAYPCQVQVQPAIPPIIMMPQ 120

DB 61 FTATVNHCHGPPLEVTQVTPFWAKVWVWPPKASHINZAYPCQVQVQPAIPPIIMMPQ 120

QY 61 FTATVNHCHGPPLEVTQVTPFWAKVWVWPPKASHINZAYPCQVQVQPAIPPIIMMPQ 120

DB 61 FTATVNHCHGPPLEVTQVTPFWAKVWVWPPKASHINZAYPCQVQVQPAIPPIIMMPQ 120

QY 121 QASFLPATLTMTVDKGNVNIISFKKVLIKEDAVIYKNGSFTHSVPRHREVPDILEVHLPH 180  
 DB 121 QASFLPATLTMTVDKGNVNIISFKKVLIKEDAVIYKNGSFTHSVPRHREVPDILEVHLPH 180

QY 181 AQPQAGVYSARYIGSNIFTSAFTLIVPRFAKWKWPECHILTA-MNNWV-HEITGEV 240  
 DB 181 AQPQAGVYSARYIGSNIFTSAFTLIVPRFAKWKWPECHILTA-MNNWV-HEITGEV 240

QY 241 ICPPEFMHRTTEKAKPEHTEFTKESQSGFQKSEVFCSTATWKKGLQNE 300  
 DB 241 ICPPEFMHRTTEKAKPEHTEFTKESQSGFQKSEVFCSTATWKKGLQNE 300

QY 301 ACHPGFYGDCKLRKSCNNGHMDRFGGCLCSQWGLQEREGDIPRMIPKIVDLPHLE 460  
 DB 301 ACHPGFYGDCKLRKSCNNGHMDRFGGCLCSQWGLQEREGDIPRMIPKIVDLPHLE 460

QY 361 VNSGKFNPIKASGMPLEPTNEEMTLVPRDGTVLPRKDNHIDHFSVAITTHILPDSG 420  
 DB 361 VNSGKFNPIKASGMPLEPTNEEMTLVPRDGTVLPRKDNHIDHFSVAITTHILPDSG 420

QY 421 VVWCVSVNTVAGWVEKFNISVKVLPKELNAPVDTGHNFAVINISSEYFG 472  
 DB 421 VVWCVSVNTVAGWVEKFNISVKVLPKELNAPVDTGHNFAVINISSEYFG 472

RESULT 2  
 protein-tyrosine kinase (EC 2.7.1.112), receptor type tek precursor bovine  
 N:Alternate names: receptor tyrosine kinase tie-2  
 C:Species: Bos primigenius taurus (cattle)  
 C>Date: 28-Oct-1995 #sequence\_revision 01-Nov-1995 #text\_change 16-Jul-1999  
 C:Accession: S57846; S32691  
 R:Sato, T.N.; Qin, Y.; Kozak, C.A.; Audus, K.L.  
 Proc. Natl. Acad. Sci. U.S.A. 96: 9355-9358, 1993  
 A:Title: tie-1 and tie-2 define another class of putative receptor tyrosine kinase genes  
 A:Reference number: S57845; MUII-94023374  
 A:Accession: S57846  
 A>Status: preliminary; nucleic acid sequence not shown, translation not shown  
 A:Molecule type: mRNA  
 A:Residues: 1-1125 <SA2>  
 A:Cross-references: EMBL:X71424; MID:q296577; PDB:1AA50555.1; PDB:q296578  
 A>Note: the nucleotide sequence was submitted to the EMBL Data Library, March 1994  
 R:Sato, T.N.; Qin, Y.; Kozak, C.A.; Audus, K.L.  
 submitted to the EMBL Data Library, March 1993  
 A:Reference number: S32690  
 A:Accession: S32699  
 A:Molecule type: mRNA  
 A:Residues: 1-1125 <SA2>  
 A:Cross-references: EMBL:X71424; MID:q296577; PDB:1AA50555.1; PDB:q296578  
 C:Superfamily: Protein tyrosine kinase, receptor type tie; EGF homology; fibronectin type  
 C:Keywords: ATP, autophosphorylation, diphosphatase, glycoprotein, phosphoprotein, phospho  
 E:1-22/Domain: signal sequence #status predicted <SI>  
 E:24-104/Domain: protein tyrosine kinase, receptor type tek #status predicted <MAT>  
 E:145-137/Region: cell attachment (R-G-D) motif  
 E:211-251/Domain: EGF homology <EG1>  
 E:255-298/Domain: EGF homology <EG2>  
 E:302-340/Domain: EGF homology <EG3>  
 E:342-426/Domain: immunoglobulin homology <IM>  
 E:447-527/Domain: fibronectin type III repeat homology <FN1A>  
 E:542-626/Domain: fibronectin type III repeat homology <FN3B>  
 E:639-721/Domain: fibronectin type III repeat homology <FN3C>  
 E:753-774/Domain: transmembrane #status predicted <TM>  
 E:823-1100/Domain: protein kinase homology <KIN>  
 E:831-840/Region: protein kinase ATP-binding motif  
 E:140,158,399,438,464,560,597,650,692/Binding site: carbohydrate (Asn) (covalent) #statu  
 E:856,873,965/Active site: Lys, Glu, Asp #status predicted

Query Match: 93.58; Score 2435; DB 1; Length 1125;  
 Best Local Similarity 93.28; Pred. No. 5e-161;  
 Matches 440; Conservative 12; Mismatches 20; Indels 0; Gaps 0;

QY 1 MDSLASLVGVCSLLSGIVEGAMGLILNSLPLVSLAEISLTICIASGWPHPEPTIGRD 60  
 DB 1 MDSLASLVGVCSLLSGIVEGAMGLILNSLPLVSLAEISLTICIASGWPHPEPTIGRD 60

QY 61 FEALMNHQDPLFTVQTVTPFAWKKVVMKPEKASKINCAVYCEGSPVGAIRIPTMKMPQ 120  
 DB 61 FEALMNHQDPLFTVQTVTPFAWKKVVMKPEKASKINCAVYCEGSPVGAIRIPTMKMPQ 120

QY 121 QASFLPATLTMTVDKGNVNIISFKKVLIKEDAVIYKNGSFTHSVPRHREVPDILEVHLPH 180  
 DB 121 QASFLPATLTMTVDKGNVNIISFKKVLIKEDAVIYKNGSFTHSVPRHREVPDILEVHLPH 180

QY 181 AQPQAGVYSARYIGSNIFTSAFTLIVPRFAKWKWPECHILTA-MNNWV-HEITGEV 240  
 DB 181 AQPQAGVYSARYIGSNIFTSAFTLIVPRFAKWKWPECHILTA-MNNWV-HEITGEV 240

QY 241 ICPPEFMHRTTEKAKPEHTEFTKESQSGFQKSEVFCSTATWKKGLQNE 300  
 DB 241 ICPPEFMHRTTEKAKPEHTEFTKESQSGFQKSEVFCSTATWKKGLQNE 300

QY 301 ACHPGFYGDCKLRKSCNNGHMDRFGGCLCSQWGLQEREGDIPRMIPKIVDLPHLE 460  
 DB 301 ACHPGFYGDCKLRKSCNNGHMDRFGGCLCSQWGLQEREGDIPRMIPKIVDLPHLE 460

QY 361 VNSGKFNPIKASGMPLEPTNEEMTLVPRDGTVLPRKDNHIDHFSVAITTHILPDSG 420  
 DB 361 VNSGKFNPIKASGMPLEPTNEEMTLVPRDGTVLPRKDNHIDHFSVAITTHILPDSG 420

QY 421 VVWCVSVNTVAGWVEKFNISVKVLPKELNAPVDTGHNFAVINISSEYFG 472  
 DB 421 VVWCVSVNTVAGWVEKFNISVKVLPKELNAPVDTGHNFAVINISSEYFG 472

RESULT 3  
 154237  
 protein tyrosine kinase (EC 2.7.1.112), receptor type tek precursor mouse  
 N:Alternate names: protein-tyrosine kinase receptor type tie2  
 C:Species: Mus sp. (mouse)  
 C>Date: 02-Aug-1996 #sequence\_revision 02-Aug-1996 #text\_change 21-Jan-2000  
 C:Accession: 154237  
 R:Eutenei, A.S.; Stackel, S.A.; Wilks, A.F.  
 Growth Factors 9: 99-105, 1993  
 A:Title: tie2, a putative protein tyrosine kinase from a new class of cell surface re  
 A:Reference number: 154237; MUII-94031116  
 A:Accession: 154237  
 A>Status: preliminary; translated from GH/EMBL/DDHJ  
 A:Molecule type: mRNA  
 A:Residues: 1-1125 <PE5>  
 A:Cross-references: GB:S67051; NID:q452873; PDB:1AAB26663.1; PDB:q452874  
 C:Genetics:  
 A:Gene: tie2  
 C:Superfamily: Protein tyrosine kinase, receptor type tie; EGF homology; fibronectin  
 C:Keywords: ATP, autophosphorylation, diphosphatase, glycoprotein, phosphoprotein, phospho  
 E:27-104/Domain: immunoglobulin homology <IM>  
 E:211-251/Domain: EGF homology <EG1>  
 E:255-298/Domain: EGF homology <EG2>  
 E:302-340/Domain: EGF homology <EG3>  
 E:540-624/Domain: fibronectin type III repeat homology <3FR>  
 E:820-1097/Domain: protein kinase homology <FIN>  
 E:828-836/Region: protein kinase ATP binding motif

Query Match: 92.48; Score 2407; DB 2; Length 1125;  
 Best Local Similarity 41.98; Pred. No. 4.3e-159;  
 Matches 434; Conservative 17; Mismatches 21; Indels 0; Gaps 0;

QY 1 MDSLASLVGVCSLLSGIVEGAMGLILNSLPLVSLAEISLTICIASGWPHPEPTIGRD 60  
 DB 1 MDSLASLVGVCSLLSGIVEGAMGLILNSLPLVSLAEISLTICIASGWPHPEPTIGRD 60

QY 61 FEALMNHQDPLFTVQTVTPFAWKKVVMKPEKASKINCAVYCEGSPVGAIRIPTMKMPQ 120  
 DB 61 FEALMNHQDPLFTVQTVTPFAWKKVVMKPEKASKINCAVYCEGSPVGAIRIPTMKMPQ 120



Db 61 FEALMNGHLEPLEVLCVTRFEWAKKVVWKEPKASKINAYPPEKVPRLAIPPTMKMPQ 120  
 QY 121 QASFLPATLMTVDKGNVNISSKKVLEKELAVIYKNGSPHSHVPRHEVPIILVHLPH 180  
 Db 121 QASFLPATLMTVDKGNVNISSKKVLEKELAVIYKNGSPHSHVPRHEVPIILVHLPH 180  
 QY 181 AQPDACVYSARYIGNLPTSAFTRIIVPPEAKW-HRNNHILATAWNGWHECTDPE 240  
 Db 181 AQPDACVYSARYIGNLPTSAFTRIIVPPEAKW-HRNNHILATAWNGWHECTDPE 240  
 QY 241 ICYPCWERTCOKACMLHPTFCRCKPCKRSGQCKSYVQLPVPVQSSATCKWKLQNF 300  
 Db 241 ICYPCWERTCOKACMLHPTFCRCKPCKRSGQCKSYVQLPVPVQSSATCKWKLQNF 300  
 QY 301 ACHPGYGHACKLRSCNNDFMTRKDDGKCSQWQWQDEPEGPRMIPKIVLPDHL 360  
 Db 301 ACHPGYGHACKLRSCNNDFMTRKDDGKCSQWQWQDEPEGPRMIPKIVLPDHL 360  
 QY 361 VNSCKENPICKASGWLPTNEENTIVKPKCTVLRKQENHTHFSVALPTIIPILPPSG 420  
 Db 361 VNSCKENPICKASGWLPTNEENTIVKPKCTVLRKQENHTHFSVALPTIIPILPPSG 420  
 QY 421 VWCSVNTVAGWKEKPNISVVKLPKPLNAPVLDTGHNFAVINISSEPYFG 472  
 Db 421 VWCSVNTVAGWKEKPNISVVKLPKPLNAPVLDTGHNFAVINISSEPYFG 472

RESULT 4  
 JN0712  
 N:Alternate names: protein-tyrosine kinase, receptor type tek precursor mouse  
 C:Species: Mus musculus (house mouse)  
 C>Date: 27 Jun-1994 #sequence\_revision 27 Jun-1994 #text\_change 16-Jul-1999  
 C:Accession: JN0712; S57848; S43495; S33142  
 R:Iwama, A.; Hanaguchi, I.; Hashiyama, M.; Murayama, Y.; Yasunaga, K.; Suda, T.  
 Biochem. Biophys. Res. Commun. 195, 301-309, 1993  
 A:Title: Molecular cloning and characterization of mouse TIE and TEK receptor tyrosine kinase  
 A:Reference number: JN0711; MUID:93371421  
 A:Accession: JN0712  
 A:Molecule type: mRNA  
 A:Residues: 1-1123 <IWA>  
 R:Sato, T.N.; Qin, Y.; Kozak, C.A.; Audus, K.L.  
 Proc. Natl. Acad. Sci. U.S.A. 90, 9355-9358, 1993  
 A:Title: tie-1 and tie-2 define another class of putative receptor tyrosine kinase genes  
 A:Reference number: S57845; MUID:94022374  
 A:Accession: S57848  
 A>Status: preliminary; nucleic acid sequence not shown, translation not shown  
 A:Molecule type: mRNA  
 A:Residues: 1-786,788 1123 <SAT>  
 A:Cross-references: EMBL X71427; NID:9296612; PID:9296613  
 A:Note: the nucleotide sequence was submitted to the EMBL data library, March 1993  
 R:Dumont, D.J.  
 submitted to the EMBL Data library, April 1993  
 A:Reference number: S43495  
 A:Accession: S43495  
 A:Molecule type: mRNA  
 A:Residues: 1-537,539,735,737,737-786,788-1123 <IWA>  
 A:Cross-references: EMBL X67553; NID:9297158; PID:9297159  
 R:Dumont, D.J.; Yamaguchi, I.P.; Conlon, P.A.; Bossant, J.; Reitman, M.L.  
 Oncogene 7, 1471-1480, 1992  
 A:Title: tek, a novel tyrosine kinase gene located on mouse chromosome 4, is expressed in  
 A:Reference number: S43494; MUID:92334855  
 A:Accession: S43494  
 A:Molecule type: mRNA  
 A:Residues: 823-1123 <DUW>  
 A:Cross-references: EMBL X67553  
 C:Comment: Receptor tyrosine kinase ligand systems play an important role in the constitutive  
 C:Genetics:  
 A:Map position: 4  
 C:Superfamily: protein-tyrosine kinase, receptor type tek, EGF homology: fibronectin  
 C:Keywords: ATP, autophosphorylation, duplication, glycoprotein, phosphoprotein, phospho  
 E1-22/20main: signal sequence status predicted <SIG>  
 E:23-1123/Product: protein-tyrosine kinase, receptor type tek #status predicted <MAT>

F:12-104/Domain: immunoglobulin homology <IM>  
 F:135-137/Region: cell attachment (R-G-D) motif  
 F:211-251/Domain: EGF homology <EG1>  
 F:255-298/Domain: EGF homology <EG2>  
 F:302-312/Domain: EGF homology <EG3>  
 F:364-426/Domain: immunoglobulin homology <IM>  
 F:447-525/Domain: fibronectin type III repeat homology <FN3B>  
 F:540-624/Domain: fibronectin type III repeat homology <FN3B>  
 F:637-719/Domain: fibronectin type III repeat homology <FN3C>  
 F:751-771/Domain: transmembrane #status predicted <TM>  
 F:821-1009/Domain: protein kinase homology <PK>  
 F:829-837/Region: protein kinase ATP-binding motif  
 F:140,158,499,458,545,546/Binding site: carbohydrate (Asn) (covalent) #st  
 F:854,871,963/Active site: Lys, Glu, Asp #status predicted

Query Match 92.48; Score 2407; DB 1; Length 1123;  
 Best Local Similarity 91.98; Pred. No. 4.4e-159;  
 Matches 434; Conservative 17; Mismatches 21; Indels 0; Gaps 0;

QY 1 MSLASIVLCQVSLISSEIVECAMLILINSLPVSUAKLSLICIASOWRHPHPIICRD 60  
 Db 1 MSLASIVLCQVSLISSEIVECAMLILINSLPVSUAKLSLICIASOWRHPHPIICRD 60  
 QY 61 FEALMNGHLEPLEVLCVTRFEWAKKVVWKEPKASKINAYPPEKVPRLAIPPTMKMPQ 120  
 Db 61 FEALMNGHLEPLEVLCVTRFEWAKKVVWKEPKASKINAYPPEKVPRLAIPPTMKMPQ 120  
 QY 121 QASFLPATLMTVDKGNVNISSKKVLEKELAVIYKNGSPHSHVPRHEVPIILVHLPH 180  
 Db 121 QASFLPATLMTVDKGNVNISSKKVLEKELAVIYKNGSPHSHVPRHEVPIILVHLPH 180  
 QY 181 AQPDACVYSARYIGNLPTSAFTRIIVPPEAKW-HRNNHILATAWNGWHECTDPE 240  
 Db 181 AQPDACVYSARYIGNLPTSAFTRIIVPPEAKW-HRNNHILATAWNGWHECTDPE 240  
 QY 241 ICYPCWERTCOKACMLHPTFCRCKPCKRSGQCKSYVQLPVPVQSSATCKWKLQNF 300  
 Db 241 ICYPCWERTCOKACMLHPTFCRCKPCKRSGQCKSYVQLPVPVQSSATCKWKLQNF 300  
 QY 301 ACHPGYGHACKLRSCNNDFMTRKDDGKCSQWQWQDEPEGPRMIPKIVLPDHL 360  
 Db 301 ACHPGYGHACKLRSCNNDFMTRKDDGKCSQWQWQDEPEGPRMIPKIVLPDHL 360  
 QY 361 VNSCKENPICKASGWLPTNEENTIVKPKCTVLRKQENHTHFSVALPTIIPILPPSG 420  
 Db 361 VNSCKENPICKASGWLPTNEENTIVKPKCTVLRKQENHTHFSVALPTIIPILPPSG 420  
 QY 421 VWCSVNTVAGWKEKPNISVVKLPKPLNAPVLDTGHNFAVINISSEPYFG 472  
 Db 421 VWCSVNTVAGWKEKPNISVVKLPKPLNAPVLDTGHNFAVINISSEPYFG 472

RESULT 5  
 JN0771  
 C:protein-tyrosine kinase (E1-23-1123), receptor type tek precursor mouse  
 C:Species: Mus musculus (house mouse)  
 C>Date: 13 Jun-1993 #sequence\_revision 10 Jun-1993 #text\_change 21-Jul-2000  
 C:Accession: JN0771  
 R:Horita, K.; Yagi, T.; Kohmura, N.; Tomooka, Y.; Ikawa, Y.; Aizawa, S.  
 Biochem. Biophys. Res. Commun. 189, 1747-1753, 1992  
 A:Title: A novel tyrosine kinase, tek, expressed in murine embryonic stem cells.  
 A:Reference number: JN0771; MUID:93129253  
 A:Accession: JN0771  
 A:Molecule type: mRNA  
 A:Residues: 1-1125 <HP>  
 A:Cross-references: GR P13738; NID:9220439; PID:9220439  
 C:Genetics:  
 A:Gene: tek  
 C:Superfamily: protein-tyrosine kinase, receptor type tek, EGF homology: fibronectin  
 C:Keywords: ATP, autophosphorylation, duplication, glycoprotein, phosphoprotein, phospho  
 E1-22/20main: signal sequence status predicted <SIG>  
 E:23-1125/Product: protein-tyrosine kinase, receptor type tek #status predicted <MAT>

F:47-104/Domain: immunoglobulin homology <IM1>  
 F:135-137/Region: cell attachment (P-6.6) motif  
 F:212-252/Domain: EGF homology <E21>  
 F:256-299/Domain: EGF homology <E22>  
 F:303-341/Domain: EGF homology <E23>  
 F:365-427/Domain: immunoglobulin homology <IM2>  
 F:448-526/Domain: fibronectin type III repeat homology <FN1A>  
 F:541-625/Domain: fibronectin type III repeat homology <FN1B>  
 F:638-720/Domain: fibronectin type III repeat homology <FN1C>  
 F:752-773/Domain: transmembrane status predicted <TM>  
 F:824-1100/Domain: protein kinase homology <KIN>  
 F:831-849/Region: protein kinase ATP-binding motif  
 F:140-158,400,439,465,559,649,691/Binding site: carbohydrate (Asn) (covalent) \*status  
 F:856,873,965/Active site: Lys, Glu, Asp \*status predicted

Query Match 89.7% Score 2337.5 DB 1 Length 1125,  
 Best Local Similarity 89.9% Pred. No. 2,9e-154  
 Matches 426, Conserved 16, Mismatches 26, Indels 3, Gaps 2

QY 1 MDSIASLVGVSLLSGTVHGMHLLINSPLVSDAFSLTCTIASGWPDPHPPTIGRD 60  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 DB 1 MDSIAGLVGVSLLSGTVHGMHLLINSPLVSDAFSLTCTIASGWPDPHPPTIGRD 60  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 QY 61 PEALNQIDPLEVTGVTTFWAKKVVWKKPEKAKINATYFTSTPVSRREAITPTWKMPG 120  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 DB 61 PEALMNQIDPLEVTGVTTFWAKKVVWKKPEKAKINATYFTSTPVSRREAITPTWKMPG 120  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 QY 121 QASFLPATITWVKGDNVNISSPKKVLIKEDAVIKNGSFIR--SVPRHVPDILLEVH 178  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 DB 121 QASFLPATITWVKGDNVNISSPKKVLIKEDAVIKNGS--LEPLSAPGMKYLIDLEVIL 179  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 QY 179 PHAQDQACVYSAYIGCNLFTSAPTRIVWRPEAKWGPCNHLCTACMNNGVCHEDTG 238  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 DB 180 PHAQDQACVYSAYIGCNLFTSAPTRIVWRPEAKWGPCNHLCTACMNNGVCHEDTG 239  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 QY 239 EITTEFTEMHPTTEKATLHTFPTTEKPSQVWVDFDFYGSCTATGKGLQNE 298  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 DB 240 EITTEFTEMHPTTEKATLHTFPTTEKPSQVWVDFDFYGSCTATGKGLQNE 299  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 QY 299 PHAQDQACVYSAYIGCNLFTSAPTRIVWRPEAKWGPCNHLCTACMNNGVCHEDTG 358  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 DB 300 NEAPPSGYGPDCKLRCHTNEETIDREFGDCSSGWSGLQCEKGRKWTPTIETLPPH 359  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 QY 359 LEVNSGKFNPLCKASGWLPTNEMTLVKPDVILHPKDNHDSVAIFTHRIILPPD 418  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 DB 360 LEVNSGKFNPLCKASGWLPTNEMTLVKPDVILHPKDNHDSVAIFTHRIILPPD 419  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 QY 419 SGVWVGSVNTVAGVMEKPFENISVKVLKPLNAPNVIIDIRHNAVITSSPPYFG 472  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 DB 420 SGVWVGSVNTVAGVMEKPFENISVKVLKPLNAPNVIIDIRHNAVITSSPPYFG 473  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 6  
 S57845  
 protein-tyrosine kinase (EC 2.7.1.112), receptor type tie precursor - bovine  
 C:Species: Bos primigenius taurus (cattle)  
 C>Date: 28-Oct-1995 #sequence, revision 03-Nov-1995 #text change 16-Jul-1999  
 C:Accession: S57845; S3640  
 R:Sato, T.N.; Qin, Y.; Kozak, C.A.; Audus, K.L.  
 Proc. Natl. Acad. Sci. U.S.A. 90 9456-9458 1993  
 A:Title: Tie-1 and Tie-2 define another class of putative receptor tyrosine kinase genes  
 A:Reference number: S57845; MUID:94022374  
 A:Accession: S57845  
 A:Status: preliminary, nucleic acid sequence not shown, translation not shown  
 A:Molecule type: mRNA  
 A:Residues: 1-1136 <SAT>  
 A:Cross-references: EMRL: X71424; NID:12966575; PIRN:AA0554 1; PFD:429667  
 A:Note: The nucleotide sequence was submitted to the EMBL Data Library, March 1993  
 C:Superfamily: protein tyrosine kinase, receptor type tie, EGF homology, fibronectin type  
 C:Keywords: ATP, autophosphorylation, duplication, glycoprotein, phosphoprotein, phospho  
 F:1-21/Domain: signal sequence \*status predicted <SIG>  
 F:22-1146/Region: protein tyrosine kinase, receptor type tie \*status predicted <MAT>

F:336-108/Domain: immunoglobulin homology <IM1>  
 F:213-253/Domain: EGF homology <E21>  
 F:257-303/Domain: EGF homology <E22>  
 F:304-342/Domain: EGF homology <E23>  
 F:363-426/Domain: immunoglobulin homology <IM2>  
 F:447-528/Domain: fibronectin type III repeat homology <FN1A>  
 F:540-631/Domain: fibronectin type III repeat homology <FN1B>  
 F:640-738/Domain: fibronectin type III repeat homology <FN1C>  
 F:759-784/Domain: transmembrane status predicted <TM>  
 F:835-1112/Domain: protein kinase homology <KIN>  
 F:843-851/Region: protein kinase ATP-binding motif  
 F:43-106,376-424/Disulfide bonds: \*status predicted  
 F:84,150,501,594,707/Binding site: carbohydrate (Asn) (covalent) \*status predicted  
 F:868,885,977/Active site: Lys, Glu, Asp \*status predicted

Query Match 33.5% Score 873.5 DB 1 Length 1146,  
 Best Local Similarity 39.3% Pred. No. 9.4e-53  
 Matches 176, Conserved 16, Mismatches 176, Indels 3, Gaps 12

QY 16 TSGTVEGAMDLILINSPLVSDAFSLTCTI-----ASGWRPHEPTIIGHDFEALNN 66  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 DB 15 LASHVGAADVLTLLATLLELPDPEFLTCVSCACACRGSDAWCP--PILLEKIDKIVPT 72  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 QY 67 QH--QPLEVTQDVTPEWAKKVVWKKPEKAKINATYFTSTPVSRREAITPTWKMPG 123  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 DB 73 PRWQPP-----HIAKNGSSRVTVRGSQPSDILGVFSCVG---GGTRVLYVHNSINGAH 124  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 QY 124 FILPATITWVKGDNVNISSPKKVLIKEDAVIKNGSFTHSVPRHVPDILLEVHILPP 180  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 DB 125 LLPDKVTHTVNGDGTAVLSARVKEKQTDVWKSNGSYFTIEDHAEQDQGLQ--EPR 182  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 QY 181 APODQACVYSAYIGCNLFTSAPTRIVWRPEAKWGPCNHLCTACMNNGVCHEDTG 240  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 DB 183 VDPSSGTSYATYFASNIQSAPFPLVPRCEACGPGQCTKCTVCCICGCVCHQDQGLQ 242  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 QY 241 APTPEFTMHPPTTEKATLHTFPTTEKPSQVWVDFDFYGSCTATGKGLQNE 300  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 DB 243 VPRPPEFTSIRCEQACPPGRFGQSSQVQCPGISGCRGLIFLPPYGCSSGWSRSCQQR 302  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 QY 301 ACHPEKADQWCKLWSSNNGRMGRTFQGLTSSQGLQCEPREGIDPRMTPKIVLVDHLE 360  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 DB 303 AAPTPEFTMHPPTTEKATLHTFPTTEKPSQVWVDFDFYGSCTATGKGLQNE 359  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 QY 361 VNSGKFNPLCKASGWLPTNEMTLVKPDVILHPKDNHDSVAIFTHRIILPPD 418  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 DB 460 FNLDIMPRINCAACAGNPPVWGSMEILKKPDCTVILSTKAIVEPDR-TTAREFVPRILAGD 418  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 QY 419 SGVWVGSVNTVAGVMEKPFENISVKVLKPLNAPNVI 454  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 DB 419 SGLWPCRVSTSGGDSRRPRINVKVDPVPLIADPLL 454  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 7  
 S24066  
 protein-tyrosine kinase (EC 2.7.1.112), receptor type tie precursor - human  
 C:Species: Homo sapiens (man)  
 C>Date: 19-Feb-1994 #sequence, revision 11-Aug-1995 #text change 21-Jul-2000  
 C:Accession: S24066; C38269; I52613  
 R:Partanen, J.; Armstrong, E.; Maekelae, T.P.; Korhonen, J.; Sandberg, M.; Penkonen,  
 Mol. Cell Biol. 12, 1698-1707, 1992  
 A:Title: A novel endothelial cell surface receptor tyrosine kinase with extracellular  
 A:Reference number: S24066; MUID:92195316  
 A:Accession: S24066  
 A:Molecule type: mRNA  
 A:Residues: 1-1138 <PAR>  
 A:Cross-references: EMBL: X60957  
 R:Partanen, J.; Maekelae, T.P.; Allitalo, E.; Lehto-Saarelake, H.; Allitalo, K.  
 Proc. Natl. Acad. Sci. U.S.A. 87, 8913-8917, 1990  
 A:Title: Putative tyrosine kinases expressed in K-562 human leukemia cells.  
 A:Reference number: A38268; MUID:91062389  
 A:Accession: C38269  
 A:Status: preliminary, nucleic acid sequence not shown, not compared with conceptual

A:Molecule type: mRNA  
 A:Residues: 981-1034 <PAW>  
 A:Experimental source: clone JTK14  
 R:Korhonen, J.; Lahtinen, I.; Halmekyto, M.; Alhonen, L.; Jänne, J.; Dumont, D.; Allitala  
 Blood 86, 1828-1835, 1995  
 A:Title: Endothelial specific gene expression directed by the tie gene promoter in vivo.  
 A:Reference number: 152613, MUID:95383653  
 A:Accession: 152613  
 A:Status: translated from GR/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-19 <RES>  
 A:Cross-references: GR:573447; NID:91086920; PIDN:AAU4299.1; PID:94261999  
 C:Genetics:  
 A:Gene: GHR:TIE; JTK14  
 A:Cross-references: GRP 212973, OMIM:600222  
 A:Map position: lp34-1p33  
 C:Function:  
 A:Description: Catalyzes the phosphorylation of a peptidyl tyrosine residue by ATP  
 C:Superfamily: Protein-tyrosine kinase, receptor type tie, EGF homology, fibronectin type  
 C:Keywords: ATP, autophosphorylation, duplication, glycoprotein, phosphoprotein, phospho  
 F:1-21/Domain: signal sequence status predicted <SIG>  
 F:22-1138/Domain: protein tyrosine kinase, receptor type tie #status predicted <MAT>  
 F:36-107/Domain: immunoglobulin homology <IM1>  
 F:36-107/Domain: EGF homology <EG1>  
 F:215-255/Domain: EGF homology <EG2>  
 F:259-302/Domain: EGF homology <EG3>  
 F:306-344/Domain: EGF homology <EG4>  
 F:365-428/Domain: immunoglobulin homology <IM2>  
 F:449-530/Domain: fibronectin type III repeat homology <FNVA>  
 F:542-633/Domain: fibronectin type III repeat homology <FNVB>  
 F:642-730/Domain: fibronectin type III repeat homology <FNVC>  
 F:761-786/Domain: transmembrane #status predicted <TM>  
 F:837-1114/Domain: protein kinase homology <KIN>  
 F:845-853/Region: protein kinase ATP-binding motif  
 F:43-105,372-426/Disulfide bonds: #status predicted  
 F:83,161,503,596,709/Binding site: carbohydrate (Asn) (covalent) #status predicted  
 F:870,887,979/Active site: Lys, Glu, Asp #status predicted

Query Match 33.2%, Score 865.5, DB 1, Length 1138;  
 Best Local Similarity 38.2%, Pred. No. 3.4e 52;  
 Matches 174, Conservative 71, Mismatches 185, Indels 25, Gaps 4;  
 QY 14 LLLSGTVEGAMDLILNLSPLVNSVTSITLCI-----ASQGFPHETITIGDPEAL 64  
 DB 13 LFLASHVGAAGVGLLLANLELTLPQFFLTGVSGASVNSQAWP---PLLEKLDIV 70  
 QY 65 MNHQDPLEVTQVDTREKAKVVKWR--EKANKINAYPTGKVPVGBAIPKTKMPQAS 123  
 DB 71 FTFPTETLRF----IAPNCSHVLESEFETSSGVSEVSGAGAKKIPVIVVNSICAR 126  
 QY 124 FLPATLTMTVDKGNVNISEKKVLIKEEDNAVYKNGSFTHSVPRIEVDP---LLEVILPH 180  
 DB 127 LLDKVTHTVTKGDTAVLSARVUIKROTIVKNSGSYPTLDMIEAQQGRELQ--LPN 184  
 QY 181 AQPDAGVSARYIGGNIPTSAFTPIVPPCPACKKCPNHCITACMNGVCHDITGEC 240  
 DB 185 VQPPSSGIYSATVLEASPLASAFELIVRGNGPWPSPVTKEDPGLHGWHCHIDPEY 244  
 QY 241 ICPEPGWGRTEKACETHTFGPTKEPSSQELKSKYVFLEDPYPSVATPWKLQYNE 400  
 DB 245 VQPPGPTHTHTGACAPGCPGPGQSCQCPGQSCQCPGLPGLDNYGSCSCSGWQSCQ 404  
 QY 301 AHPREYGPVPEKLPKSNNGCMTQRFQVPLSHQWQILQTEREIPPMPTKIVELPCHIE 400  
 DB 305 AAFAPHFAPLRQLQVNNQVSLTPSPGVVPSWIDVILTEKSG---PFPQILNMASELE 401  
 QY 361 VNSGKNPI--CKASGWLPIPIEFMTLVKPDGTVLRHPKDFNHTDHFSAVITFTHRIIPDS 419  
 DB 362 FNLTPRINCAAGNFPVRSIELRKPDGTVLLSTKAIVEPEKTTATBEFVPRVLVADS 421  
 QY 420 GVWVCVNTVAGVVEKPFENLSVKVLPKPLNAPWV 454  
 DB 422 GPWECRVSTGGQDSRRPKVNVKVPVPLAAPRL 456

## RESULT 8

JN0711  
 Protein-tyrosine kinase (EC 2.7.1.112), receptor type tie precursor mouse  
 C:Species: Mus musculus (house mouse)  
 C>Date: 14 Jul 1994 #sequence\_revision 14 Jul 1994 #text\_change 16-Jul-1999  
 C:Accession: JN0711; S33141; S57847; A48926; 165403  
 R:Iwama, A.; Hamaguchi, I.; Hashiyama, M.; Mura, Y.; Yasunaga, K.; Suda, T.  
 Biochem. Biophys. Res Commun 195 301-309, 1993  
 A:Title: Molecular cloning and characterization of mouse TIE and TEK receptor tyrosin  
 A:Reference number: JN0711, MUID:93371421  
 A:Accession: JN0711  
 A:Molecule type: mRNA  
 A:Residues: 1-1134 <TWA>  
 A:Cross-references: GR:X79066; NID:9402601; PIDN:CAA52148.1; PID:9402602  
 R:Sato, T.N.; Oin, Y.; Kozak, C.A.; Andus, K.L.  
 submitted to the EMBL data library, March 1993  
 A:Reference number: S32690  
 A:Accession: S33141  
 A:Molecule type: mRNA  
 A:Residues: 1-598, 1-1500-1134 <SAT>  
 A:Cross-references: EMBL:X731425; NID:949610; PIDN:CAA50556.1; PID:9496611  
 R:Sato, T.N.; Oin, Y.; Kozak, C.A.; Andus, K.L.  
 Proc. Natl. Acad. Sci. U.S.A. 90, 9355-9358, 1993  
 A:Title: tie 1 and tie 2 define another class of putative receptor tyrosine kinase ge  
 Blood 80, 2548-2555, 1992  
 A:Reference number: S57845, MUID:94023374  
 A:Accession: S57847  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: mRNA  
 A:Residues: 1-598, 1-1500-1134 <SAT>  
 A:Cross-references: EMBL:X71425, NID:9496610, PIDN:CAA50556.1, PID:9496611  
 A:Note: the nucleotide sequence was submitted to the EMBL Data Library, March 1993  
 R:Korhonen, J.; Partanen, J.; Armstrong, E.; Vahdokari, A.; Elenius, K.; Jalkanen, M  
 Blood 80, 2548-2555, 1992  
 A:Title: Enhanced expression of the tie receptor tyrosine kinase in endothelial cells  
 A:Reference number: A48926, MUID:93043301  
 A:Accession: A48926  
 A:Status: preliminary; not compared with conceptual translation  
 A:Molecule type: mRNA  
 A:Residues: 221-527,740 899 <KRG>  
 A:Note: sequence extracted from NCI backbone (NCBIP:118660, NCBIP:118662)  
 R:Korhonen, J.; Lahtinen, I.; Halmekyto, M.; Alhonen, L.; Jänne, J.; Dumont, D.; Alit  
 Blood 86, 1828-1835, 1995  
 A:Title: Endothelial-specific gene expression directed by the tie gene promoter in vi  
 A:Reference number: 152613; MUID:95383653  
 A:Accession: 155403  
 A:Status: preliminary; translated from GR/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-19 <RES>  
 A:Cross-references: GR:S79346; NID:91086920  
 C:Superfamily: protein-tyrosine kinase, receptor type tie; EGF homology; fibronectin  
 C:Keywords: ATP, autophosphorylation, duplication, glycoprotein, phosphoprotein; phos  
 F:1-21/Domain: signal sequence status predicted <SIG>  
 F:22-1134/Domain: protein tyrosine kinase, receptor type tie #status predicted <MAT>  
 F:36-105/Domain: immunoglobulin homology <IM1>  
 F:214-254/Domain: EGF homology <EG1>  
 F:257-300/Domain: EGF homology <EG2>  
 F:304-342/Domain: EGF homology <EG3>  
 F:363-426/Domain: immunoglobulin homology <IM2>  
 F:447-528/Domain: fibronectin type III repeat homology <FNVA>  
 F:540-629/Domain: fibronectin type III repeat homology <FNVB>  
 F:648-736/Domain: fibronectin type III repeat homology <FNVC>  
 F:757-782/Domain: transmembrane #status predicted <TM>  
 F:44-110/Domain: protein kinase homology <KIN>  
 F:841-849/Region: protein kinase ATP-binding motif  
 F:43-103,370-424/Disulfide bonds: #status predicted  
 F:81,159,501,592,705/Binding site: carbohydrate (Asn) (covalent) #status predicted  
 F:866,883,975/Active site: Lys, Glu, Asp #status predicted

Query Match 31.5%, Score 820.5, DB 1, Length 1134;  
 Best Local Similarity 36.0%, Pred. No. 4.5e-49;





QY 402 CHFGYGFEDCKLRSON NEMDQDFQGLSPGQWGLGCFER EGIPTMTK 351  
 DB 552 COVGVEGDLCKP-CNCIDMTEGGWCQCTCTGCKCKKCFACDCKDCTDIAPYGYPCCKAC 610  
 QY 452 IVLDDHIFVNSGKFNPKCKASOWPLIPNEEM 384  
 DB 611 ACD-----GAGTTSPECEALISQCEPCNGNFI 636

RESULT 15  
 S18253  
 Laminin alpha-1 chain precursor - fruit fly (Drosophila melanogaster)  
 C:Species: Drosophila melanogaster  
 C>Date: 16-Sep-1992 #sequence\_revision 24-Jul-1997 #text\_change 10-Dec-1999  
 C:Accession: S28395; S18253  
 R:Kusilek-Gallberg, M.; Garrison, K.; MacKrell, A.J.; Fessler, J.H.  
 EMBL J. 11, 4519-4527, 1992  
 A>Title: Laminin A chain: expression during Drosophila development and genomic sequence.  
 A:Reference number: S28399; MUID:93049203  
 A:Accession: S28395  
 A>Status: preliminary  
 A:Molecule type: nucleic acid  
 A:Residues: 1-4712 <KUS>  
 A:Cross-references: GR-906498; NID-q157799; PIDN:AAA28662.1; PID:q157800  
 R:Garrison, K.; MacKrell, A.J.; Fessler, J.H.  
 J. Biol. Chem. 266, 22899-22904, 1991  
 A>Title: Drosophila laminin A chain sequence, interspecies comparison, and domain structure  
 A:Reference number: S18253; MUID:92078147  
 A:Accession: S18253  
 A:Molecule type: mRNA  
 A:Residues: 1762-3712 <GAR>  
 A:Cross-references: EMBL:M75882; NID:q157797; PIDN:AAA28661.1; PID:q157798  
 C:Genetics:  
 A:Gene: FlyBase: FBar.0002526  
 A:Cross-references: FlyBase: FBar.0002526  
 C:Superfamily: laminin alpha-1 chain; laminin G repeat homology; laminin-type EGF-like  
 C:Keywords: basement membrane; cell binding; coiled coil; disulfide bond; extracellular  
 F:273-340/Domain: laminin-type EGF-like homology <LE3>  
 F:333-400/Domain: laminin-type EGF-like homology <LE2>  
 F:541-584/Domain: laminin-type EGF-like homology <LE02>  
 F:1776-2115/Domain: III <DOM3>  
 F:1776-1806/Domain: laminin-type EGF-like homology #status atypical <LE1>  
 F:1809-1856/Domain: laminin-type EGF-like homology <LE2>  
 F:1859-1914/Domain: laminin-type EGF-like homology <LE3>  
 F:1917-1967/Domain: laminin-type EGF-like homology <LE4>  
 F:1970-2014/Domain: laminin-type EGF-like homology <LE5>  
 F:2017-2061/Domain: laminin-type EGF-like homology <LE6>  
 F:2044-2109/Domain: laminin-type EGF-like homology <LE7>  
 F:2116-2697/Domain: I/II, heptad repeats <DOM2>  
 F:2698-3712/Domain: G <DOM6>  
 F:2698-2863/Domain: repeat G1 <RG1>  
 F:2864-3048/Domain: repeat G2 <RG2>  
 F:3049-3223/Domain: repeat G3 <RG3>  
 F:3079-3260/Domain: laminin G repeat homology <LG3>  
 F:3334-3528/Domain: repeat G4 <RG4>  
 F:3529-3712/Domain: repeat G5 <RG5>  
 F:1847,1850,1943,2024,2196,2215,2267,2301,2123,2492,2524,2538,2569,2699,2720,2890,2938,3

Query Match 8.1%; Score 212; DB 2; Length 3712;  
 Best Local Similarity 24.6%; Pred. No. 2.5e-06;  
 Matches 70; Conservative 38; Mismatches 87; Indels 90; Gaps 18;

QY 210 RCEAKWG PRCNHLCTACMNGVCHDCEGICDPCGFCMRTCKKACELHTFCRTCKRC 268  
 DB 1407 KCRSYVGFDPCKP--CKCPNSAMCEPTTGECPMPNVIGDLCEK-CAPTYG-----F 1457  
 QY 269 SQCEKCKSVFCLDPQYV-----GSCATGKMKGLQCNFACHGFGV-PDCKL 313  
 DB 1458 HOVIGCEE--CACNPMCIANGNSOCHIFNGICHCQNIFGRAC-IVCSNGVFNPHCE- 1512  
 QY 314 RGSNNG---EMCDREUG-CLCSFHWGLQCE-----REGIP----- 346

EP 1513 QCSHKEP:LELFV-LKILH-A-FYKKNVWGPQCGVGTNYLLESNPDGTTTFCFCKETS 1472  
 QY 347 -----RMTPKIVDLPDHLFVNSGKFN-PICKASOWPLIPNEEM---ILKPKKIVLHPKD 397  
 DB 1573 RCD SAYLHVYVWVSLKHKVSTITTFEHEESKFKDMWVYADLILNFIILKADFTLEVND 1632  
 QY 398 -----FNUTDHFVS---AIFTHI --RILPPD 418  
 DB 1633 EPPAYFGVLDYLLNGNNHISAYGGHAYTLHFISCPDGKYIVAD 1677

Search completed: September 23, 2002, 23:15:22  
 Job time: 138 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd

OM protein - protein search, using sw model

Run on: September 23, 2002, 22:15:24, Search time 17.75 seconds  
(without alignments)  
1029 613 Million cell updates/sec

Title: US-09-733-764-2\_copy\_1\_472

Perfect score: 2605

Sequence: 1 MDSIASLVICGVSLIIISGV.....VIDTGHNFVNISSEYFG 472

Scoring table: BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 2030000000

Post processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: SwissProt\_40\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result No	Score	Query		Length	DP	ID	Description
		Match	%				
1	2505	100	0	1124	1	TIE2_HUMAN	Q03763 homo sapien
2	2435	93.5	1	1125	1	TIE2_BOVIN	Q06807 bos taurus
3	2407	92.4	1	1122	1	TIE2_MOUSE	Q02858 mus musculus
4	873.5	33.5	1	1136	1	TIE1_BOVIN	Q06805 bos taurus
5	865.5	33.2	1	1138	1	TIE1_HUMAN	P35590 homo sapien
6	820.5	31.5	1	1134	1	TIE1_MOUSE	Q06806 mus musculu
7	289	11.1	830	1	SREC_HUMAN	Q14162 homo sapien	
8	231.5	8.9	685	1	DLL4_HUMAN	Q9NF61 homo sapien	
9	221	8.5	3672	1	LMC2_CAEEL	Q21313 caenorhabdi	
10	212	8.1	3712	1	LMADROME	Q00174 drosophila	
11	206.5	7.9	1964	1	NTC4_MOUSE	P41895 mus musculus	
12	195.5	7.7	686	1	GLI4_MOUSE	Q6J171 mus musculu	
13	195.5	7.7	1408	1	SERR_DROME	P18168 drosophila	
14	198.5	7.6	379	1	WIF1_HUMAN	Q9Y5W5 homo sapien	
15	198.5	7.6	833	1	DL1_DROME	P10041 drosophila	
16	198	7.5	723	1	DLL1_HUMAN	Q03548 homo sapien	
17	195.5	7.5	379	1	WIF1_MOUSE	Q06807 mus musculu	
18	193	7.4	2444	1	NTC1_HUMAN	P46531 homo sapien	
19	192.5	7.4	722	1	DLL1_MOUSE	Q61483 mus musculu	
20	192.5	7.4	1429	1	LI12_CAEEL	P14585 caenorhabdi	
21	192	7.4	2635	1	LMAS5_HUMAN	Q15230 homo sapien	
22	191.5	7.4	2524	1	NOTC_XENLA	P21783 xenopus lae	
23	191.5	7.4	3718	1	LMAS5_MOUSE	Q61001 mus musculu	
24	190.5	7.3	714	1	DLL1_RAT	P97677 rattus norv	
25	189	7.3	473	1	FP2_MYTGA	Q25464 mytilus gal	
26	187.5	7.2	2318	1	NOTC3_MOUSE	Q61982 mus musculu	
27	185.5	7.1	2703	1	NOTC_DROME	F07207 drosophila	
28	184.5	7.1	1808	1	TENA_CHICK	P10039 gallus gali	
29	184	7.1	1295	1	GLP1_CAEEL	P13598 caenorhabdi	
30	183	7.0	2481	1	UN52_CAEEL	Q06561 caenorhabdi	
31	183	7.0	2531	1	NTC1_MOUSE	Q01705 mus musculu	
32	182	7.0	2531	1	NTC1_RAT	Q07008 rattus norv	
33	181.5	7.0	2437	1	NOTC_HHARE	P46530 brachydanio	

# RESULTS

TIE2\_HUMAN  
ID TIE2\_HUMAN  
AC Q02763  
STANDARD: PRT: 1124 AA.

DT 01-FEB-1994 (Rel. 28, Created)

DT 01-FEB-1994 (Rel. 28, Last sequence update)

DI 16-DEC-2001 (Rel. 43, Last annotation update)

DE Angiopoietin 1 receptor precursor (EC 2.7.1.112) (Tyrosine-protein

kinase receptor Tib.2) (Tyrosine-protein kinase receptor TEK) (P140

DE TEK) (Tunica interna endothelial cell kinase).

GN TEK OR TIE2.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.

OX NCBI\_TaxID:9606;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Placenta;

RX MEDLINE=93173509; PubMed=8382358;

RA Tiegler S.F., Bird T.A., Schottinger J.A., Schooley K.A., Haum P.K.;

RT "Molecular cloning and characterization of a novel receptor protein

tyrosine kinase from human placenta."

RL Oncogene 8:663-670(1993).

RN [2]

RP VARIANT VCMC1 TRP-849.

RX MEDLINE=9716065; PubMed=8780625;

RA Vikkula M., Boon L.M., Carraway K.L. III, Calvert J.T., Diamanti A.J.;

RA Goumarov B., Pasyk K.A., Marchuk D.A., Warman M.L., Cantley L.C.;

RA Mulliken J.B., Olse R.P.;

RT "Vascular dysmorphogenesis caused by an activating mutation in the

receptor tyrosine kinase TIE2."

RL Cell 87:1181-1190(1996).

RN [3]

RP VARIANTS VCMC1 TRP-849 AND SER-897.

RX MEDLINE=9729245; PubMed=1053874;

RA Calvert J.T., Riney T.T., Kontos C.D., Cha F.H., Prieto V.G.;

RA Shea C.R., Berg J.N., Nevin N.C., Simpson S.A., Pasyk K.A.;

RA Speer M.C., Peters K.G., Marchuk D.A.;

RT "Allelic and locus heterogeneity in inherited venous malformations."

RL Hum Mol Gene 8:1279-1289(1999).

CC -!- FUNCTION: THIS PROTEIN IS A PROTEIN TYROSINE-KINASE TRANSMEMBRANE

RECEPTOR FOR ANGIOPOIETIN 1. IT MAY CONSTITUTE THE EARLIEST

MAMMALIAN ENDOTHELIAL CELL LINEAGE MARKER. PROBABLY REGULATES

ENDOTHELIAL CELL PROLIFERATION, DIFFERENTIATION AND GUIDES THE

PROPER PATTERNING OF ENDOTHELIAL CELLS DURING BLOOD VESSEL

FORMATION.

CC -!- CATALYTIC ACTIVITY: ATP + a protein tyrosine - ADP + protein

tyrosine phosphate.

CC -!- SUBCELLULAR LOCATION: Type I membrane protein.

CC -!- TISSUE SPECIFICITY: PREDOMINANTLY EXPRESSED IN ENDOTHELIAL CELLS

AND THEIR PROGENITORS. THE ANGIOBLASTS HAS BEEN DIFFICULTY FOUND

IN PLACENTA AND LUNG, WITH A LOWER LEVEL IN UMBILICAL VEIN

ENDOTHELIAL CELLS, SPRAIN AND KIDNEY.

CC -!- DISEASE: DEFECTS IN TEK ARE A CAUSE OF DOMINANTLY INHERITED VENOUS

MALFORMATIONS (VCMC1), AN ERROR OF VASCULAR MORPHOGENESIS

CHARACTERIZED BY DILATED, SERPIGINOUS CHANNELS.

## ALIGNMENTS

34 179 6.9 1064 1 FBPI\_STRPU  
35 178.5 6.9 2201 1 TFNA\_HUMAN  
36 178 6.8 592 1 DLL3\_MOUSE  
37 178 6.8 3110 1 LMA2\_HUMAN  
38 176 6.8 2139 1 CFB\_DROME  
39 174 6.7 577 1 ITH6\_CAVPO  
40 174 6.7 618 1 DLL3\_HUMAN  
41 173.5 6.7 589 1 DLL3\_RAT  
42 173.5 6.7 1376 1 CRBL\_HUMAN  
43 172.5 6.7 1159 1 LMR1\_MOUSE  
44 172.5 6.6 1798 1 LMR2\_HUMAN  
45 172 6.6 1746 1 TENA\_FIC  
P10079 strongyloce  
P24921 homo sapien  
Q88216 mus musculu  
P24043 homo sapien  
P10040 drosophila  
P18563 cavia porce  
Q9NY17 homo sapien  
Q88671 rattus norv  
P82279 homo sapien  
Q51987 mus musculu  
P55268 homo sapien  
Q29116 sus scrofa

CC -1- SIMILARITY: TO OTHER PROTEIN-TYROSINE KINASES IN THE CATALYTIC  
 CC DOMAIN.  
 CC -1- SIMILARITY: CONTAINS 2 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.  
 CC -1- SIMILARITY: CONTAINS 3 FIBRONECTIN TYPE III-LIKE DOMAINS.  
 CC -1- SIMILARITY: CONTAINS 3 EGF-LIKE DOMAINS.

CC THIS SWISS-PROT entry is copyrighted. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcements/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).

CC -----  
 CC EMBL: L06139; AAA61139.1;  
 CC HSP: P11462; IFGK.  
 CC MIM: 600221;  
 CC MIM: 600195;  
 CC InterPro: IPR000561; EGF-like.  
 CC InterPro: IPR000719; FcR-kinase.  
 CC InterPro: IPR003961; FN-III.  
 CC InterPro: IPR001245; Tyr\_kinase.  
 CC Pfam: PF00008; EGF\_1.  
 CC Pfam: PF00041; fn3\_3.  
 CC Pfam: PF00069; kinase\_1.  
 CC PRINTS: PR00109; TYRKINASE.  
 CC SMART: SM00181; EGF\_2.  
 CC SMART: SM00001; EGF-like\_1.  
 CC SMART: SM00060; FN3\_3.  
 CC SMART: SM00219; Tyrc\_1.  
 CC PROSITE: PS00107; PROTEIN\_KINASE\_ATP\_1.  
 CC PROSITE: PS00109; PROTEIN\_KINASE\_TYR\_1.  
 CC PROSITE: PS00011; PROTEIN\_KINASE\_DOM\_1.  
 CC PROSITE: PS00022; EGF\_1\_3.  
 CC PROSITE: PS01186; EGF\_2\_3.  
 CC Receptor; Tyrosine-protein kinase; Transferase; Signal; ATP-binding;  
 CC Repeat; EGF-like domain; Transmembrane; Immunoglobulin domain;  
 CC Glycoprotein; phosphorylation; Multigene family; Disease mutation.

CC SIGNAL 1 18 POTENTIAL.  
 CC CHAIN 19 1124 ANGIOPOIETIN 1 RECEPTOR.  
 CC DOMAIN 19 745 EXTRACELLULAR (POTENTIAL).  
 CC TRANSFM 746 770 POTENTIAL.  
 CC DOMAIN 771 1124 CYTOPLASMIC (POTENTIAL).  
 CC DOMAIN 44 102 IC-LIKE C2-TYPE DOMAIN 1.  
 CC DOMAIN 210 252 EGF-LIKE 1.  
 CC DOMAIN 254 299 EGF-LIKE 2.  
 CC DOMAIN 301 341 EGF-LIKE 3.  
 CC DOMAIN 370 424 IC-LIKE C2-TYPE DOMAIN 2.  
 CC DOMAIN 444 536 FIBRONECTIN TYPE-III 1.  
 CC DOMAIN 541 634 FIBRONECTIN TYPE-III 2.  
 CC DOMAIN 638 732 FIBRONECTIN TYPE-III 3.  
 CC DOMAIN 824 1096 PROTEIN KINASE.  
 CC NP\_BIND 830 838 ATP (BY SIMILARITY).  
 CC BINDING 855 855 ATP (BY SIMILARITY).  
 CC ACT\_SITE 964 964 BY SIMILARITY.  
 CC CARBOHYD 140 140 N-LINKED (GLCNAC...) (POTENTIAL).  
 CC CARBOHYD 158 158 N-LINKED (GLCNAC...) (POTENTIAL).  
 CC CARBOHYD 399 399 N-LINKED (GLCNAC...) (POTENTIAL).  
 CC CARBOHYD 438 438 N-LINKED (GLCNAC...) (POTENTIAL).  
 CC CARBOHYD 464 464 N-LINKED (GLCNAC...) (POTENTIAL).  
 CC CARBOHYD 560 560 N-LINKED (GLCNAC...) (POTENTIAL).  
 CC CARBOHYD 596 596 N-LINKED (GLCNAC...) (POTENTIAL).  
 CC CARBOHYD 649 649 N-LINKED (GLCNAC...) (POTENTIAL).  
 CC CARBOHYD 691 691 N-LINKED (GLCNAC...) (POTENTIAL).  
 CC MOD\_RES 942 942 PHOSPHORYLATION (AUTO-).  
 CC VARIANT 849 849 R -> W (IN VMCM1; ACTIVATING EFFECT).  
 CC VARIANT 897 897 Y -> S (IN VMCM1; ACTIVATING EFFECT).  
 CC SEQUENCE 1124 AA; 125810 MW; 65805018FA40C6C664;

Query Match

100.00% Score 2605; DR 1; Length 1124;

Best local similarity 100.0%; Pred. No. 9,70-200;  
 Matches 472; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MDLSLVLCGVSLILSGVSGVAMDLILINSLPLVSDAETSLTCLASWRPHEPTIGRD 60  
 DB 1 MDLSLVLCGVSLILSGVSGVAMDLILINSLPLVSDAETSLTCLASWRPHEPTIGRD 60  
 QY 61 FEALMNHQDPLEVTDVTFEWAQKVVWKKKASKINAYFEGVPRVGAIRIKMKMRQ 120  
 DB 61 FEALMNHQDPLEVTDVTFEWAQKVVWKKKASKINAYFEGVPRVGAIRIKMKMRQ 120  
 QY 121 QASFI PATITMI VIKGINVINSPKKVI IKEPDAVIYKNGSFTHSVPRHPKVPDILVHLPH 180  
 DB 121 QASFI PATITMI VIKGINVINSPKKVI IKEPDAVIYKNGSFTHSVPRHPKVPDILVHLPH 180  
 QY 181 AOPDAGVYSARYIGGNLFTSAFIRLIVRCEAOKWGPCNHLCTACNNGVCHEDTGR 240  
 DB 181 AOPDAGVYSARYIGGNLFTSAFIRLIVRCEAOKWGPCNHLCTACNNGVCHEDTGR 240  
 QY 241 IGRPGMGRICEKACELIRFGRICNEKSGQCKSVFCLPDPYCCSCATGKGLQNE 300  
 DB 241 IGRPGMGRICEKACELIRFGRICNEKSGQCKSVFCLPDPYCCSCATGKGLQNE 300  
 QY 301 ACHPGYGPCKIKRCSNNGPMCDRFGGICSGWGGICQFREGIPRMKIPKIVDLPHIE 360  
 DB 301 ACHPGYGPCKIKRCSNNGPMCDRFGGICSGWGGICQFREGIPRMKIPKIVDLPHIE 360  
 QY 361 VNSGKNPICKASGWLPTNEEMTLVKPDGTVLHDKDFHUTDHSVAITTHRLIPUSG 420  
 DB 361 VNSGKNPICKASGWLPTNEEMTLVKPDGTVLHDKDFHUTDHSVAITTHRLIPUSG 420  
 QY 421 VVCSVNTVAGMYEKPFNISVKVLPKPLNAPNVIETGHINFAVINISSEPYFG 472  
 DB 421 VVCSVNTVAGMYEKPFNISVKVLPKPLNAPNVIETGHINFAVINISSEPYFG 472

## RESULT 2

11E2\_BOVIN  
 ID TIE2\_BOVIN STANDARD: PRT: 1125 AA.  
 AC Q06807;  
 DT 01-FEB-1995 (Rel. 31, Created)  
 DT 01-FEB-1995 (Rel. 31, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Angiopoietin 1 receptor precursor (EC 2.7.1.112) (Tyrosine-protein  
 DE kinase receptor TIE-2).  
 GN TEK OR TIE2 OR TIE-2.  
 OS Bos taurus (Bovine).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 OC Bovidae; Bovinae; Bos.  
 OX NCBI\_TaxID:9913;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-Endothelial cells;  
 RX MEDLINE=9402374; PubMed=8415706;  
 PA Sato T.N., Gira Y., Kozak C.A., Andus K.L.;  
 FT Tie-1 and Tie-2 define another class of putative receptor tyrosine  
 FT kinase genes expressed in early embryonic vascular system.\*;  
 FL Proc. Natl. Acad. Sci. U.S.A. 90:9355-9358(1993).  
 CC -1- FUNCTION: THIS PROTEIN IS A PROTEIN-TYROSINE-KINASE TRANSMEMBRANE  
 CC RECEPTOR FOR ANGIOPOIETIN 1. IT MAY CONSTITUTE THE EARLIEST  
 CC MAMMALIAN ENDOTHelial CELL LINEAGE MARKER. PROBABLY REGULATES  
 CC ENDOTHelial CELL PROLIFERATION, DIFFERENTIATION AND GUIDES THE  
 CC PROPER PATTERNING OF ENDOTHelial CELLS DURING BLOOD VESSEL  
 CC FORMATION.  
 CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine -> ADP + protein  
 CC tyrosine phosphate.  
 CC -1- SUBCELLULAR LOCATION: type i membrane protein.  
 CC -1- TISSUE SPECIFICITY: SPECIFICALLY EXPRESSED IN DEVELOPING VASCULAR  
 CC ENDOTHelial CELLS.  
 CC -1- SIMILARITY: TO OTHER PROTEIN-TYROSINE KINASES IN THE CATALYTIC  
 CC DOMAIN.  
 CC -1- SIMILARITY: CONTAINS 2 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.



CC -1- SIMILARITY: CONTAINS 3 FIBRONECTIN TYPE III-LIKE DOMAINS.  
 CC -1- SIMILARITY: CONTAINS 3 EGF-LIKE DOMAINS.  
 CC This SWISS-PROT entry is copyright, it is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: X71424; CAA50555.1; -  
 DR PIR: S42691; S32691.  
 DR ISSP: P11362; IFGR.  
 DR InterPro: IPR003561; EGF-like.  
 DR InterPro: IPR00719; Euk pk kinase.  
 DR InterPro: IPR003961; FN.II  
 DR InterPro: IPR002049; Laminin\_EGF.  
 DR InterPro: IPR001245; Tyl\_pkinase.  
 DR Pfam: PF00008; EGF; 2.  
 DR Pfam: PF00041; fn3; 3.  
 DR Pfam: PF00069; pkinase; 1.  
 DR PRINTS: PR00109; IYRKINASE.  
 DR SMART: SM00180; EGF\_Lam; 1.  
 DR SMART: SM00001; EGF\_Like; 1.  
 DR SMART: SM00060; FN3; 3.  
 DR SMART: SM00219; Tyrc; 1.  
 DR PROSITE: PS00107; PROTEIN\_KINASE\_AIP; 1.  
 DR PROSITE: PS00109; PROTEIN\_KINASE\_TYR; 1.  
 DR PROSITE: PS00011; PROTEIN\_KINASE\_DOM; 1.  
 DR PROSITE: PS00022; EGF\_1; 3.  
 DR PROSITE: PS01186; EGF\_2; 2.  
 KW receptor, tyrosine protein kinase, transferase, signal, ATP-binding;  
 KW Repeat, EGF-like domain, Transmembrane, Immunoglobulin domain,  
 KW Glycoprotein, Phosphorylation, Multigene family.  
 FT SIGNAL 1 18  
 FT CHAIN 19 1125  
 FT DOMAIN 19 746  
 FT TRANSMEM 747 771  
 FT POTENTIAL.  
 FT CYTOPLASMIC (POTENTIAL).  
 FT IC-LIKE C2-TYPE DOMAIN 1  
 FT EGF-LIKE 1.  
 FT EGF-TYPE 2.  
 FT EGF-LIKE 3.  
 FT IC-LIKE C2-TYPE DOMAIN 2.  
 FT FIBRONECTIN TYPE-III 1.  
 FT FIBRONECTIN TYPE-III 2.  
 FT FIBRONECTIN TYPE-III 3.  
 FT PROTEIN KINASE.  
 FT NP\_BIND 831 839  
 FT BINDING 856 856  
 FT ATP (BY SIMILARITY).  
 FT HY SIMILARITY.  
 FT ACT\_SITE 965 965  
 FT MOD\_RES 993 993  
 FT CARBOHYD 158 158  
 FT SEQUENCE 1125 AA; 125/27 MW; 015F1320AB853B7F CRC64;  
 Query March 93 58; Score 2435, DR 1, Length 1125,  
 Best local Similarity 93.2%; Pred. No. 3, 4e-186;  
 Matches 440; Conservative 12; Mismatches 20; Indels 0; Gaps 0;  
 QY 1 MDSLASLVLCGVSLLSGTVEAMDLILNSPLVSDAETSLTCTASWPPPEPTITIGPD 60  
 DB 1 MDSLACVLCGVSLLSAIVDCAMDLILNSPLVSDAETSLTCTASWPPPEPTITIGPD 60  
 QY 61 FEALMCHQDPLEVTADVTPEWAKKVVWPEKASKINAYPPFVPPEAIPITMKMPQ 120  
 DB 61 FEALMCHQDPLEVTQDATPEWAKKVVWPEKASKINAYFCEG3VPPEAIPITMKMPQ 120  
 QY 121 QASFLPATITMTVDKGVNVSFKKVLKPEDAVLYKNSFIRSHVPRHVPVDILEVHLPH 180  
 DB 121 QASFLPATITMTVDKGVNVSFKKVLKPEDAVLYKNSFIRSHVPRHVPVDILEVHLPH 180

QY 181 AQPDAQVYSARYTGGNLTSAFTSLVPRPEAKWKEPENNHTATAMNNVWVEHTSEC 240  
 DB 181 AQPDAQVYSARYTGGNLTSAFTSLVPRPEAKWKEPENNHTATAMNNVWVEHTSEC 240  
 QY 241 ICPPPEFMKICEKACAEHLFKICKERCSQDQCKSVFCLPDPYQSCAIGWKGLOQNE 300  
 DB 241 ICPPPEFMKICEKACAEHLFKICKERCSQDQCKSVFCLPDPYQSCAIGWKGLOQNE 300  
 QY 301 ACPQGVYGVCKKICWSSNNHMTLRFQWGLCSQWGLQCKEKGELPQWPKVPLPHIE 360  
 DB 301 ACPQGVYGVCKKICWSSNNHMTLRFQWGLCSQWGLQCKEKGELPQWPKVPLPHIE 360  
 QY 361 VNSCKPNDPCASQWPLTREEMTLVPRDGTVLHKPKFNHTDMPFSAITTHILPDSG 420  
 DB 361 VNSCKPNDPCASQWPLTREEMTLVPRDGTVLHKPKFNHTDMPFSAITTHILPDSG 420  
 QY 421 VVWLSVNTVACMVEKPENISVKVLKPLNAPNVDTCNFAV:NISSEPYG 472  
 DB 421 VVWLSVNTVACMVEKPENISVKVLKPLNAPNVDTCNFAV:NISSEPYG 472  
 RESULT 3  
 TIE2\_MOUSE  
 ID TIE2\_MOUSE STANDARD: PRT: 1122 AA.  
 AC Q02858; 1994 (Rel. 28, Created)  
 DT 01-FEB-1995 (Rel. 31, last sequence update)  
 DI 16-OCT-2001 (Rel. 40, last annotation update)  
 DE Angiopoietin 1 receptor precursor (ic 2.7.1.12) (Tyrosine-protein  
 DE kinase receptor TIE-2) (Tyrosine-protein kinase receptor TEK) (P140  
 DE TEK) (Tunica interna endothelial cell kinase) (HYK).  
 GN TEK or TIE2 or TIE-2 or HYK  
 OS Mus musculus (Mouse).  
 OC Mammalia; Eutheria; Chordata, Craniata, Vertebrata, Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BAIR/C; TISSUE=Lung;  
 RA MEDLINE=94022374; PubMed=8415706;  
 PT Sato T.N., Qin Y., Kozak G.A., Andus K.H.,  
 PT "Tie-1 and tie-2 define another class of putative receptor tyrosine  
 PT kinase genes expressed in early embryonic vascular system.";   
 PL Proc Natl Acad Sci U S A 90:9355-9358(1993).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CD-1; TISSUE=Embryonic heart;  
 RA MEDLINE=93241731; PubMed=8386827;  
 PT Dumont D.T., Gradwohl G.T., Feq G.H., Auerbach R., Breitman M.L.;  
 PT "The endothelial-specific receptor tyrosine kinase, tek, is a member  
 PT of a new subfamily of receptors.";   
 RN Oncogene 8:1293-1301(1993).  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Embryonic stem cells;  
 RA MEDLINE=93129253; PubMed=1283811;  
 PT Horita K., Yagi T., Kohmura N., Tomooka Y., Ikawa Y., Aizawa S.;  
 PT "A novel tyrosine kinase, hyk, expressed in murine embryonic stem  
 PT cells.";   
 PL Biochem Biophys Res Commun 189:1747-1753(1992).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Lung;  
 RA MEDLINE=94031116; PubMed=8217221;  
 PT Kuntz A.S., Stacker S.A., Wilks A.P.;  
 PT "Tie2, a putative protein tyrosine kinase from a new class of cell  
 PT surface receptor.";   
 PL Growth Factors 9:99-105(1993).  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RA MEDLINE=94244487; PubMed=8187650;  
 RA Schuerch H., Risau W.;  
 PT "Expression of tie-2, a member of a novel family of receptor tyrosine

kinases, in the endothelial cell lineage.\*;  
Development 119:957-968(1993).

[6].

SEQUENCE OF 822-1122 FROM N.A.

STRAIN-CD-1; JUSSON-Embryonic heart;

MEDLINE-9243485; PubMed-1630810.

RA Dumont D.J., Yamauchi T.P., Coulon R.A., Rossant J., Breitman M.L.;

\*Tek, a novel tyrosine kinase gene located on mouse chromosome 4, is

expressed in endothelial cells and their presumptive precursors.\*;

Proc Natl Acad Sci USA 90:1471-1480(1992).

CC -1- FUNCTION: THIS PROTEIN IS A PROTEIN TYROSINE-KINASE TRANSMEMBRANE

RECEPTOR FOR ANGIOPOIETIN 1. IT MAY CONSTITUTE THE EARLIEST

MAMMALIAN ENDOTHELIAL CELL LINEAGE MARKER. PROBABLY REGULATES

ENDOTHELIAL CELL PROLIFERATION, DIFFERENTIATION AND SHIPES THE

PROPER PATTERNING OF ENDOTHELIAL CELLS DURING BLOOD VESSEL

FORMATION.

CC -1- CATALYTIC ACTIVITY: ATP + a protein-tyrosine -> ADP + protein

tyrosine phosphate.

CC -1- SUBCELLULAR LOCATION: Type I membrane protein.

CC -1- TISSUE SPECIFICITY: SPECIFICALLY EXPRESSED IN DEVELOPING VASCULAR

ENDOTHELIAL CELLS.

CC -1- DEVELOPMENTAL STAGE: EXPRESSION DETECTABLE IN DAY 8.5 EMBRYOS

CC -1- SIMILARITY: TO OTHER PROTEIN-TYROSINE KINASES IN THE CATALYTIC

DOMAIN.

CC -1- SIMILARITY: CONTAINS 2 IMMUNOGLULIN-LIKE C2-TYPE DOMAINS

CC -1- SIMILARITY: CONTAINS 3 FIBRONECTIN TYPE III-LIKE DOMAINS.

CC -1- SIMILARITY: CONTAINS 3 EGF-LIKE DOMAINS.

CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration

between the Swiss Institute of Bioinformatics and the EMBL outstation -

the European Bioinformatics Institute. There are no restrictions on its

use by non profit institutions as long as its content is in no way

modified and this statement is not removed. Usage by and for commercial

entities requires a license agreement (See <http://www.isb-sib.ch/annotation/>

or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).

CC EMBL: X71426; CAA50557.1;

CC EMBL: X67553; CAA42857.1;

CC EMBL: D13738; BAA02883.1;

CC EMBL: S67051; AAB28663.1;

CC PIR: S33142; S33142.

CC HSSP: P11362; IFPK.

CC MGI: 98664; Tek.

CC InterPro: IPR000561; EGF-like

CC InterPro: IPR000719; Euk\_pkinase.

CC InterPro: IPR001961; FNIII.

CC InterPro: IPR001245; Tyr\_kinase.

CC Pfam: PF00041; In3; 3.

CC Pfam: PF00069; pkinase; 1.

CC PRINTS: PR00109; TYRKINASE.

CC SMART: SM00181; EGF; 2.

CC SMART: SM00060; FN3; 2.

CC SMART: SM00219; Tyrc; 1.

CC PROSITE: PS00107; PROTEIN KINASE ATP; 1.

CC PROSITE: PS00109; PROTEIN KINASE TYR; 1.

CC PROSITE: PS00111; PROTEIN KINASE DOM; 1.

CC PROSITE: PS00122; EGF\_1; 3.

CC PROSITE: PS01186; EGF\_2; 3.

CC Receptor: Tyrosine protein kinase; transferase, signal; APP-binding;

CC Repeat: EGF-like domain; Transmembrane, Immunoglobulin domain;

CC Glycoprotein; Phosphorylation; Multigene family.

CC SIGNAL: 1 18 POTENTIAL.

CC CHAIN: 19 1122 ANGIOPOIETIN 1 RECEPTOR.

CC DOMAIN: 19 744 EXTRACELLULAR (POTENTIAL).

CC TRANSMEM: 745 769 POTENTIAL.

CC DOMAIN: 779 1122 CYTOPLASMIC (POTENTIAL).

CC MEDIAN ENDOTHELIAL CELL LINEAGE MARKER. PROBABLY REGULATES

ENDOTHELIAL CELL PROLIFERATION, DIFFERENTIATION AND SHIPES THE

PROPER PATTERNING OF ENDOTHELIAL CELLS DURING BLOOD VESSEL

FORMATION.

CC -1- CATALYTIC ACTIVITY: ATP + a protein-tyrosine -> ADP + protein

tyrosine phosphate.

CC -1- SUBCELLULAR LOCATION: Type I membrane protein.

CC -1- TISSUE SPECIFICITY: SPECIFICALLY EXPRESSED IN DEVELOPING VASCULAR

ENDOTHELIAL CELLS.

CC -1- DEVELOPMENTAL STAGE: EXPRESSION DETECTABLE IN DAY 8.5 EMBRYOS

CC -1- SIMILARITY: TO OTHER PROTEIN-TYROSINE KINASES IN THE CATALYTIC

DOMAIN.

CC -1- SIMILARITY: CONTAINS 2 IMMUNOGLULIN-LIKE C2-TYPE DOMAINS

CC -1- SIMILARITY: CONTAINS 3 FIBRONECTIN TYPE III-LIKE DOMAINS.

CC -1- SIMILARITY: CONTAINS 3 EGF-LIKE DOMAINS.

CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration

between the Swiss Institute of Bioinformatics and the EMBL outstation -

the European Bioinformatics Institute. There are no restrictions on its

use by non profit institutions as long as its content is in no way

modified and this statement is not removed. Usage by and for commercial

entities requires a license agreement (See <http://www.isb-sib.ch/annotation/>

or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).

CC EMBL: X71426; CAA50557.1;

CC EMBL: X67553; CAA42857.1;

CC EMBL: D13738; BAA02883.1;

CC EMBL: S67051; AAB28663.1;

CC PIR: S33142; S33142.

CC HSSP: P11362; IFPK.

CC MGI: 98664; Tek.

CC InterPro: IPR000561; EGF-like

CC InterPro: IPR000719; Euk\_pkinase.

CC InterPro: IPR001961; FNIII.

CC InterPro: IPR001245; Tyr\_kinase.

CC Pfam: PF00041; In3; 3.

CC Pfam: PF00069; pkinase; 1.

CC PRINTS: PR00109; TYRKINASE.

CC SMART: SM00181; EGF; 2.

CC SMART: SM00060; FN3; 2.

CC SMART: SM00219; Tyrc; 1.

CC PROSITE: PS00107; PROTEIN KINASE ATP; 1.

CC PROSITE: PS00109; PROTEIN KINASE TYR; 1.

CC PROSITE: PS00111; PROTEIN KINASE DOM; 1.

CC PROSITE: PS00122; EGF\_1; 3.

CC PROSITE: PS01186; EGF\_2; 3.

CC Receptor: Tyrosine protein kinase; transferase, signal; APP-binding;

CC Repeat: EGF-like domain; Transmembrane, Immunoglobulin domain;

CC Glycoprotein; Phosphorylation; Multigene family.

CC SIGNAL: 1 18 POTENTIAL.

CC CHAIN: 19 1122 ANGIOPOIETIN 1 RECEPTOR.

CC DOMAIN: 19 744 EXTRACELLULAR (POTENTIAL).

CC TRANSMEM: 745 769 POTENTIAL.

CC DOMAIN: 779 1122 CYTOPLASMIC (POTENTIAL).

CC MEDIAN ENDOTHELIAL CELL LINEAGE MARKER. PROBABLY REGULATES

ENDOTHELIAL CELL PROLIFERATION, DIFFERENTIATION AND SHIPES THE

PROPER PATTERNING OF ENDOTHELIAL CELLS DURING BLOOD VESSEL

FORMATION.

CC -1- CATALYTIC ACTIVITY: ATP + a protein-tyrosine -> ADP + protein

tyrosine phosphate.

CC -1- SUBCELLULAR LOCATION: Type I membrane protein.

CC -1- TISSUE SPECIFICITY: SPECIFICALLY EXPRESSED IN DEVELOPING VASCULAR

ENDOTHELIAL CELLS.

CC -1- DEVELOPMENTAL STAGE: EXPRESSION DETECTABLE IN DAY 8.5 EMBRYOS

CC -1- SIMILARITY: TO OTHER PROTEIN-TYROSINE KINASES IN THE CATALYTIC

DOMAIN.

FT DOMAIN 637 731 FIBRONECTIN TYPE-III 3.  
FT DOMAIN 822 1094 PROTEIN KINASE  
FT NP\_BIND 828 836 ATP (BY SIMILARITY).  
FT BINDING 853 853 ATP (BY SIMILARITY).  
FT ACT\_SITE 962 962 BY SIMILARITY.  
FT CARBOHYD 146 146 N-LINKED (GLUCNA...) (POTENTIAL).  
FT CARBOHYD 158 158 N-LINKED (GLUCNA...) (POTENTIAL).  
FT CARBOHYD 399 399 N-LINKED (GLUCNA...) (POTENTIAL).  
FT CARBOHYD 438 438 N-LINKED (GLUCNA...) (POTENTIAL).  
FT CARBOHYD 464 464 N-LINKED (GLUCNA...) (POTENTIAL).  
FT CARBOHYD 558 558 N-LINKED (GLUCNA...) (POTENTIAL).  
FT CARBOHYD 595 595 N-LINKED (GLUCNA...) (POTENTIAL).  
FT CARBOHYD 648 648 N-LINKED (GLUCNA...) (POTENTIAL).  
FT CARBOHYD 690 690 N-LINKED (GLUCNA...) (POTENTIAL).  
FT MOD\_RES 990 990 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).  
FT CONFLICT 161 171 FHSVPRHVP -> LHPLSAUGMKYL (IN REF. 3).  
FT CONFLICT 538 538 S -> G (IN REF. 1 AND 4).  
FT CONFLICT 736 736 A -> G (IN REF. 1 AND 4).  
FT CONFLICT 745 761 MLLIALGSAAGTCITV -> DATIHHPVWVWDEASPC (IN REF. 3).  
FT CONFLICT 786 786 N -> NV (IN REF. 3).  
FT CONFLICT 913 913 R -> G (IN REF. 3).  
FT CONFLICT 925 931 ALANSTA -> CHROOYS (IN REF. 3).  
FT CONFLICT 1117 1117 S -> P (IN REF. 3).  
SQ SEQUENCE 1122 AA: 125700 MW: 18796230.104 kDa PFE96 GRC64;

Query Match 92.4%; Score 2407; DB 1; Length 1122;  
Best Local Similarity 91.9%; Prod No. 5.7c-184;  
Matches 434; Conservative 17; Mismatches 21; Indels 0; Gaps 0;  
QY 1 MDSLASLVAGVSLISSTVERAMDTILNSTPLVSLAEISLTICIASGRPHIEPTIGRD 60  
DB 1 MSLIAGVLCVSVIHYGVVEGAMDTILNSTPLVSLAEISLTICIASGRPHIEPTIGRD 60  
QY 61 FEALMNOHQDPLEVTQDVTREWAKKVVWKEKASKINGAYFCEGVPVGAIPRTMKMPQ 120  
DB 61 FEALMNOHQDPLEVTQDVTREWAKKVVWKEKASKINGAYFCEGVPVGAIPRTMKMPQ 120  
QY 121 QASPLPATLIMTVDKGNVNISEKVKLIKEDDAVIYKNGSFTHSVPRHVEVPDILEVHLPH 180  
DB 121 QASPLPATLIMTVDKGNVNISEKVKLIKEDDAVIYKNGSFTHSVPRHVEVPDILEVHLPH 180  
QY 181 AQPDACVYVYAPYVGNFTSAPTRITVPPCEAKKWDPENELCTACMANVYCHETDEP 240  
DB 181 AQPDACVYVYAPYVGNFTSAPTRITVPPCEAKKWDPENELCTACMANVYCHETDEP 240  
QY 241 ICPPGEMGRTCEKACPLHTGRTCKECSQBERKSVVFLPDPYGSATCKWGLQONE 300  
DB 241 ICPPGEMGRTCEKACPLHTGRTCKECSQBERKSVVFLPDPYGSATCKWGLQONE 300  
QY 301 ACHPGYGPPIWKIACSNNGEMCTHFGGDIQSPGWGDIQCPHFGIIPPMIPKIVDLPDHP 360  
DB 301 ACHPGYGPPIWKIACSNNGEMCTHFGGDIQSPGWGDIQCPHFGIIPPMIPKIVDLPDHP 360  
QY 361 VNSCKFNPTCKASQWPEPTNFEMTLVPPCTVLPPKCFNHLHPSVAISLHLPLPPSSG 420  
DB 361 VNSCKFNPTCKASQWPEPTNFEMTLVPPCTVLPPKCFNHLHPSVAISLHLPLPPSSG 420  
QY 421 VWVSVNTVAGVMEKPEPNISVKVLPKPLNAPNVDTGHNFAVINISSEPYFG 472  
DB 421 VWVSVNTVAGVMEKPEPNISVKVLPKPLNAPNVDTGHNFAVINISSEPYFG 472

RESULT 4

TIE1\_BOVIN

ID TIE1\_BOVIN

AC Q36805

DT 01-FEB-1995 (Rel. 31, Created)

DI 01-FEB-1995 (Rel. 41, Last sequence update)

DE 15-JUL-1999 (Rel. 38, Last annotation update)

GN TIE1 OR TIE-1

STANDARD: PRT: 1136 AA.

DT 01-FEB-1995 (Rel. 31, Created)

DI 01-FEB-1995 (Rel. 41, Last sequence update)

DE 15-JUL-1999 (Rel. 38, Last annotation update)

GN TIE1 OR TIE-1

OS Bos taurus (Bovine).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 OC Bovinae; Bovinae; Bos.  
 OX NCBI\_TaxID=9913;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Endothelial cells;  
 RA MEDLINE=94022374, PubMed=8415706,  
 RX Sato T N., Qin Y., Kozak C A., Andus K L.  
 RI "Tie-1 and tie-2 define another class of putative receptor tyrosine  
 RT kinase genes expressed in early embryonic vascular system.";  
 RI Proc Natl Acad Sci U S A. 90:9355-9358(1993).  
 CC 1- FUNCTION: PREPARE: PROTEIN-TYROSINE-KINASE: TRANSMEMBRANE RECEPTOR.  
 CC 1- CATALYTIC ACTIVITY: ATP + a protein tyrosine + ADP + protein  
 CC tyrosine phosphate.  
 CC 1- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC 1- TISSUE SPECIFICITY: SPECIFICALLY EXPRESSED IN DEVELOPING VASCULAR  
 CC ENDOTHELIAL CELLS.  
 CC 1- SIMILARITY: TO OTHER PROTEIN-TYROSINE KINASES IN THE CATALYTIC  
 CC DOMAIN.  
 CC 1- SIMILARITY: CONTAINS 2 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.  
 CC 1- SIMILARITY: CONTAINS 3 FIBRONECTIN TYPE III-LIKE DOMAINS.  
 CC 1- SIMILARITY: CONTAINS 3 EGF-LIKE DOMAINS.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL: X71423; CAA50554.1; -  
 CC PIR: S35690; S32670.  
 DR HSP; P11362; LFGR.  
 DR InterPro: IPR003561; EGF-like.  
 DR InterPro: IPR000719; Euk\_Pkinase.  
 DR InterPro: IPR003961; FN-III.  
 DR InterPro: IPR003599; Ig.  
 DR InterPro: IPR003006; Ig\_MHC.  
 DR InterPro: IPR001245; Tyr\_Pkinase.  
 DR Pfam: PF00008; EGF; 2.  
 DR Pfam: PF00041; fn3; 3.  
 DR Pfam: PF00047; Ig; 2.  
 DR Pfam: PF00069; pkinase; 1.  
 DR PRINTS: PR00109; TYRKINASE.  
 DR SMART: SM00181; EGF; 2.  
 DR SMART: SM00060; FN3; 2.  
 DR SMART: SM00409; FN3; 1.  
 DR SMART: SM00219; Tyrc; 1.  
 DR PROSITE: PS00107; PROTEIN\_KINASE\_ATP; 1.  
 DR PROSITE: PS00104; PROTEIN\_KINASE\_TYP; 1.  
 DR PROSITE: PS00011; PROTEIN\_KINASE\_DOM; 1.  
 DR PROSITE: PS00022; EGF\_1; 3.  
 DR PROSITE: PS01186; EGF\_2; 3.  
 KW Receptor; Tyrosine protein kinase; Transferase; Signal; ATP-binding;  
 KW Repeat; EGF-like domain; Transmembrane; Immunoglobulin domain;  
 KW Glycoprotein; Phosphorylation; Multigene family.  
 FT SIGNAL: 1 23  
 FT CHAIN: 24 1136  
 FT DOMAIN: 24 757  
 FT TRANSMEM: 758 782  
 FT DOMAIN: 783 1136  
 FT DOMAIN: 43 106  
 FT DOMAIN: 212 254  
 FT DOMAIN: 256 301  
 FT DOMAIN: 303 343  
 FT DOMAIN: 370 424  
 FT DOMAIN: 444 538  
 FT DOMAIN: 541 637  
 FT DOMAIN: 641 742  
 FT DOMAIN: 837 1116  
 FT PROTEIN KINASE.

FT NP\_BIND 843 851 ATP (BY SIMILARITY).  
 FT BINDING 858 868 ATP (BY SIMILARITY).  
 FT ACT\_SITE 977 977 BY SIMILARITY.  
 FT MOD\_RES 1005 1005 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).  
 FT CARBOHYD 84 84 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHYD 159 159 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHYD 501 501 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHYD 594 594 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHYD 707 707 N-LINKED (GLCNAC...) (POTENTIAL).  
 SS SEQUENCE 1136 AA: 124953 MW: 118580.434 kDa CR664;  
 Query Match 33.5%, Score 873.5; DB 1; Length 1136;  
 Best local Similarity 39.3%, Pred. No. 8,206;  
 Matches 179; Conservative 68; Mismatches 176; Indels 33; Gaps 12;  
 QY 16 LSGTVEGAMDLILNSLFLVSDAETSLTCL -- ASGWKPHPEIIIGRDEALMN 66  
 DB 15 LASHGAAVDITLLADIRLTPQRFETCVSGEAGAGSGDAWGP--PILLEKDDRIIVRT 72  
 QY 67 QH--QDPLEVLTQVTPREWAKKVKWKPK-EKASKINAYEPDESPVPEAIPRTMKMRQAS 123  
 DB 73 PRPWQPP-----HIAPNCGSSPTVTRGSPQSDLVGVFSCVG---GGGTPELVYVINSFGAH 124  
 QY 124 PLPATLITMDVKGDVNIISPKKVLIKEDDAVIYKNGSPHSVPREHVPD---ILEVHLPH 180  
 DB 125 ILPOKVTHTVANKGDTAVLSARVPKPKQTDVITWKNNGSYETLDPHEADQVDFLLQ--LPN 182  
 QY 181 APLPGLA-WYSAPYLISNLFSAFTPLIVPRPEAKWKPE-NHLITAMNNWVJHEDTPEI 240  
 DB 183 VQPSGGIYSAIYLASPLSSAFTPLIVPRPEAKWKPEIPELKEGQGLHGVCHVCHNGRC 242  
 QY 241 TCPPFMGFTTEKATEIHTPRTPEKSGSGGKSVVVLDPDPYSGSCATKWLQONE 300  
 DB 243 VCPKFTTTRTEAWPFGSPGSGGKSVVVLDPDPYSGSCATKWLQONE 302  
 QY 301 AHTHTPHTPEKATKESNNKIMWHPPELQPSGSGGKSVVVLDPDPYSGSCATKWLQONE 350  
 DB 303 ACAPGFGGANWILGQVUNNGTDFPFGSGGKSVVVLDPDPYSGSCATKWLQONE 359  
 QY 361 VNSGKFNPI-CKASGMPLENTLVEPDGTV-IHPKDFNHTDHFSAVFTIIRLPPD 418  
 DB 360 FNIDIMPVFNAAAGNPPVPMKSMFLKAPVDFVLSIKAVFPHVW-LIAFVFPVPLAIGD 418  
 QY 419 SCVWVCVNTVAGVVEKPFENISVKVLPKPLNAPNVI 454  
 DB 419 SCLWEPVSTSGGQDSRFRFRINVKVPVPLTAPPELL 454  
 RESULT 5  
 TIE1\_HUMAN  
 ID TIE1\_HUMAN STANDARD: PRT: 1138 AA.  
 AC F35590;  
 DT 01-JUN-1994 (Rel. 29, Created)  
 DT 01-JUN-1994 (Rel. 29, Last sequence update)  
 DT 15-JUN-1999 (Rel. 38, Last annotation update)  
 DE Tyrosine protein kinase receptor TIE-1 precursor (EC 2.7.1.112).  
 GN TIE1 OR TIE.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA MEDLINE=94195316; PubMed=1312667;  
 RA Pattaran J., Armstrong E., Mackellae T.P., Korhonen J., Sandberg M.,  
 RA Reinken K., Kneutilla S., Huchner K., Allitalo K.  
 RT "A novel endothelial cell surface receptor tyrosine kinase with  
 RT extracellular epidermal growth factor homology domains.";  
 RL Mol. Cell. Biol. 12:1698-1707(1992).  
 RN [2]  
 PP PEVISTIONS.  
 RA Pattaran J.M.;

Submitted (JUL 1993) to the EMBL/GenBank/DDBJ databases  
 CC 1-1 FUNCTION: PROBABLE PROTEIN TYROSINE KINASE TRANSMEMBRANE RECEPTOR.  
 CC 1-1 CATALYTIC ACTIVITY: ATP + a protein tyrosine - ADP + protein  
 CC tyrosine phosphate.  
 CC 1-1 SUBCELLULAR LOCATION: Type I membrane protein.  
 CC 1-1 TISSUE SPECIFICITY: SPECIFICALLY EXPRESSED IN DEVELOPING VASCULAR  
 CC ENDOTHELIAL CELLS.  
 CC 1-1 SIMILARITY: TO OTHER PROTEIN-TYROSINE KINASES IN THE CATALYTIC  
 CC DOMAIN.  
 CC 1-1 SIMILARITY: CONTAINS 2 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.  
 CC 1-1 SIMILARITY: CONTAINS 3 FIBRONECTIN TYPE III-LIKE DOMAINS.  
 CC 1-1 SIMILARITY: CONTAINS 3 EGF-LIKE DOMAINS.  
 CC  
 CC THIS SWISS-PROT entry is copyrighted. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-Profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb.sib.ch/announce/>  
 CC or send an email to [license@sib-sb.ch](mailto:license@sib-sb.ch)).

EMBL: X60957; CAA43290.1; -  
 DB HSP: P11462; IF3K.  
 DB MIM: 600222; -  
 DB InterPro: IPR000561; EGF-like.  
 DB InterPro: IPR000719; Euk\_pkinase.  
 DB InterPro: IPR000661; FN\_III.  
 DB InterPro: IPR004006; Ig\_MHC.  
 DB InterPro: IPR001245; Tyr\_pkinase.  
 DB Pfam: PF00008; EGF; 2.  
 DB Pfam: PF00041; fn3; 3.  
 DB Pfam: PF00047; Ig; 2.  
 DB Pfam: PF00069; pkinase; 1.  
 DB PRINTS: PR00109; TYRKINASE.  
 DB SMART: SM00181; EGF; 2.  
 DB SMART: SM00060; FN3; 2.  
 DB SMART: SM00219; TyKc1; 3.  
 DB PROSITE: PS00107; PROTEIN\_KINASE\_ATP; 1.  
 DB PROSITE: PS0100; PROTEIN\_KINASE\_TYR; 1.  
 DB PROSITE: PS00011; PROTEIN\_KINASE\_DOM; 1.  
 DB PROSITE: PS00022; EGF\_1; 3.  
 DB PROSITE: PS01186; EGF\_2; 3.  
 DB KEGG: K01186; EGF-2; 3.  
 KW receptor; tyrosine-protein kinase; transferase, signal, ATP-binding;  
 KW repeat; EGF-like domain; Transmembrane, immunoglobulin domain;  
 KW glycoprotein; phosphorylation; Multigene family.  
 FT SIGNAL 1 24 POTENTIAL.  
 FT CHAIN 25 1138 TYROSINE-PROTEIN KINASE RECEPTOR TIE-1.  
 FT DOMAIN 25 759 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 760 784 POTENTIAL.  
 FT DOMAIN 785 1138 CYTOPLASMIC (POTENTIAL).  
 FT DOMAIN 43 105 IG-LIKE C2-TYPE DOMAIN 1.  
 FT DOMAIN 214 256 EGF-CTF 1.  
 FT DOMAIN 258 303 EGF-LIKE 2.  
 FT DOMAIN 305 345 EGF-LIKE 3.  
 FT DOMAIN 372 426 IG-LIKE C2-TYPE DOMAIN 2.  
 FT DOMAIN 446 540 FIBRONECTIN TYPE-III 1.  
 FT DOMAIN 543 639 FIBRONECTIN TYPE-III 2.  
 FT DOMAIN 643 744 FIBRONECTIN TYPE-III 3.  
 FT DOMAIN 839 1118 PROTEIN KINASE.  
 FT NP-HIND 845 853 ATP (HY SIMILARITY).  
 FT BINDING 870 870 ATP (HY SIMILARITY).  
 FT ACT\_SITE 979 979 HY SIMILARITY.  
 FT CARBOHYD 83 83 N-LINKED (GLNAC... ) (POTENTIAL).  
 FT CARBOHYD 161 161 N-LINKED (GLNAC... ) (POTENTIAL).  
 FT CARBOHYD 503 503 N-LINKED (GLNAC... ) (POTENTIAL).  
 FT CARBOHYD 596 596 N-LINKED (GLNAC... ) (POTENTIAL).  
 FT CARBOHYD 709 709 N-LINKED (GLNAC... ) (POTENTIAL).  
 FT MOD\_RES 1067 1067 PHOSPHORYLATION (ACTC... ) (HY SIMILARITY).  
 SQ SEQUENCE 1138 AA; 12589 MW; 3B42RE31676C5BA1 CRC64.

Query Match 33.2% Score 865.5; DB 1: Length 1138;  
 Best local Similarity 38.2% Pred. No. 3, 6e-61;

Matches 174, Conservative 71, Mismatches 185, Indels 25, Gaps 8;  
 QY 14 LLLSTVEGAMCLLLINSLPLVSDAETISL-----ASWEPHEPITIGRDFEAL 64  
 DB 13 LFLASHWAADVLLLANLKLIDLPQFELICVSAEAGAGSGDAWGP--PLLEKIDRIV 70  
 QY 65 MNOHQDPLEVTDVIREWAKKVVWKR--EKASKINGAYEFGFGRVGRGAIKIRIMKMHQOAS 123  
 DB 71 RTPGQPLR---LARNCSHQVTLIPAFSPKPSLVGVFSCVSGAGASPTVIVVINSFGAH 126  
 QY 124 FLPATLMTVDKGNVNIISFKVLLIKEDAVIYKNSPFIHSVPRIEVPD--ILEVHLPH 180  
 DB 127 LLLPKVTHITVNGKGTAVLSARVHKQKIWIKNSGYSYFILLHQAQGRFLLQ--LPN 184  
 QY 181 AGLGAGVAVSAPYVTCGLPFTSAFTPIVPPCEAGKNCPECNHLCTACMNNWTHETPEC 240  
 DB 185 VQPESSNITYNTHLEASPLASAFELIVRGAGVAGVHWPGATIKPFGCLDAGVTHIHDEG 244  
 QY 241 TQPPPEMERICEKACPIHIFERICEKESQDEKESVVEFIIPQVQCSFALGKGLQNE 300  
 DB 245 VCPPEFTTTRCELAAREGRERFSQGRQFGLISGRGLLPLLPYRGSGMSGWRGSGQOE 304  
 QY 361 ACHPAFYVSPKIPKPSNNSEMTPKPGVLPSPQWGLQCEPESEIPMTPIVDLPDITE 460  
 DB 305 ACAPGHGADCRLOCOCONGTCDRTSGCVPSGWHGVHCEKSD---RIPQILNMASELE 461  
 QY 361 VNSGKENFI--CKASGWELFTNEEMTLVKFGTVLHFKDNTHTDHSVAIFTIIRILPDS 419  
 DB 362 FNLETMPPIVNCMAAGNPPVPGSGTELPKPKPCTVLI--STKALVEPEKTTAEFEVPLVADS 421  
 QY 420 GVWCVSVNTVAGVVEKPFNLSVKVLKPKINAPNVI 454  
 DB 422 GFWEFRVSTSGGATSPSPPKVNVKVPVPLAAPRL 456  
 RESULT 6  
 TIE1\_MOUSE  
 ID TIE1\_MOUSE STANDARD: PRT: 1134 AA.  
 AC Q6806;  
 DT 01-FEB-1995 (Rel. 31, Created)  
 DT 01-FEB-1996 (Rel. 33, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Tyrosine protein kinase receptor TIE-1 precursor (NC 2.7.1.112).  
 GN TIE1 OR TIE-1.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID:10090;  
 RN 11  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BALB/C; TISSUE-Lung;  
 PX MEDLINE-94023374; PubMed-8415706;  
 PA Sato T N, Qin Y, Kovak C A, Andus K L;  
 RT "Tie 1 and tie-2 define another class of putative receptor tyrosine  
 RL kinase genes expressed in early embryonic vascular system.";  
 RN Proc. Natl. Acad. Sci. U.S.A. 90:9355-9358(1993).  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BALB/C; TISSUE-Bone marrow;  
 RA Krivosv A.V., Frisler M.A., Visser J.W.M., Helyavsky A.V.;  
 RN Submitted (JUN-1994) to the EMBL/GenBank/DDBJ databases.  
 RN 131  
 RP SEQUENCE FROM N.A.  
 RC STRAIN BALB/C; TISSUE-Liver;  
 PX MEDLINE-95391441; PubMed-8395848;  
 PA Iwama A., Hamauchi I., Hashiyama M., Murayama Y., Yasunaga K.,  
 RA Suda T;  
 RT "Molecular cloning and characterization of mouse TIE and TEK receptor  
 RL tyrosine kinase genes and their expression in hematopoietic stem  
 RN cells.";  
 RN Blochem. Biophys. Res. Commun. 195:301-309(1993).  
 RP SEQUENCE OF 1-19 FROM N.A.

RC STRAIN-129/SV, TISSUE: Liver.  
RA MEDLINE=95381653; PubMed=7655012;  
RA Korhonen J., Lahtinen I., Halmekyto M., Alhonen L., Janne J.,  
RA Dumont D., Alitalo K.  
RT "Endothelial-specific gene expression directed by the tie gene  
RT promoter in vivo".  
PL Blood 86:1828-1835(1995).  
CC -!- FUNCTION: PROBABLY PROTEIN TYROSINE KINASE TRANSMEMBRANE PROTEIN  
CC -!- CATALYTIC ACTIVITY: ATP + a protein tyrosine - ADP + protein  
CC tyrosine phosphate  
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.  
CC -!- TISSUE SPECIFICITY: SPECIFICALLY EXPRESSED IN DEVELOPING VASCULAR  
CC ENDOTHELIAL CELLS.  
CC -!- SIMILARITY: TO OTHER PROTEIN-TYROSINE KINASES IN THE CATALYTIC  
CC DOMAIN.  
CC -!- SIMILARITY: CONTAINS 3 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.  
CC -!- SIMILARITY: CONTAINS 3 FIBRONECTIN TYPE III-LIKE DOMAINS.  
CC -!- SIMILARITY: CONTAINS 3 EGF-LIKE DOMAINS.  
CC This SWISS-PROT entry is copyright, it is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcements/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: X71425; CAA50556.1; -  
DR EMBL: X80764; CAA56736.1; -  
DR EMBL: X73960; CAA52148.1; -  
DR EMBL: S79146; -; NOT\_ANNOTATED; CDS  
DR HSPB: P11362; IFGK.  
DR MCD: M51-94406; T1-1  
DR InterPro: IPR000561; EGF-like.  
DR InterPro: IPR000719; EGF\_LKINase  
DR InterPro: IPR003961; FN\_III.  
DR InterPro: IPR003599; Iq.  
DR InterPro: IPR003006; Iq\_MHC.  
DR InterPro: IPR004600; Iq\_Like  
DR InterPro: IPR001245; Tyr\_kinase.  
DR Pfam: PF00008; EGF; 2.  
DR Pfam: PF00041; fn3; 3.  
DR Pfam: PF00047; Iq; 2.  
DR Pfam: PF00069; pkinase; 1.  
DR PRINTS: PS00109; PROTEIN\_KINASE\_Atp; 1.  
DR PROSITE: PS00107; PROTEIN\_KINASE\_Tyr; 1.  
DR PROSITE: PS00109; PROTEIN\_KINASE\_Tyr; 1.  
DR PROSITE: PS00011; PROTEIN\_KINASE\_DOM; 1  
DR PROSITE: PS00022; EGF\_1; 1  
DR PROSITE: PS00186; EGF\_2; 3.  
DR PROSITE: PS00186; EGF\_2; 3.  
KW Receptor, Tyrosine protein kinase, Transmembrane, Signal, Atp-binding,  
KW Repeat, EGF-like domain, Transmembrane, Immunoglobulin domain,  
KW Glycoprotein, Phosphorylation, Multigene family.  
FT SIGNAL 1 22  
FT CHAIN 23 1134 TYROSINE PROTEIN KINASE RECEPTOR TIE 1.  
FT DOMAIN 23 755 EXTRACELLULAR (POTENTIAL).  
FT TRANSFEM 756 780 POTENTIAL.  
FT DOMAIN 781 1134 CYTOPLASMIC (POTENTIAL).  
FT DOMAIN 43 103 IG-LIKE C2-TYPE DOMAIN 1.  
FT DOMAIN 212 254 EGF-LIKE 1.  
FT DOMAIN 256 301 EGF-LIKE 2.  
FT DOMAIN 303 343 EGF-LIKE 3.  
FT DOMAIN 370 424 IG-LIKE C2-TYPE DOMAIN 2.  
FT DOMAIN 444 538 FIBRONECTIN TYPE-III 1.  
FT DOMAIN 541 635 FIBRONECTIN TYPE-III 2.  
FT DOMAIN 639 740 FIBRONECTIN TYPE-III 3.  
FT DOMAIN 835 1114 PROTEIN KINASE.  
FT NO\_BIND 841 849 ATP (BY SIMILARITY).

PI BINDING 806 806 AIP (BY SIMILARITY).  
FT AC\_SITE 975 975 BY SIMILARITY.  
FT MOD\_RES 1003 1003 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).  
FT CARBOHYD 81 81 N-LINKED (GLCNAC) (POTENTIAL).  
FT CARBOHYD 159 159 N-LINKED (GLCNAC) (POTENTIAL).  
FT CARBOHYD 501 501 N-LINKED (GLCNAC) (POTENTIAL).  
FT CARBOHYD 542 542 N-LINKED (GLCNAC) (POTENTIAL).  
FT CARBOHYD 705 705 N-LINKED (GLCNAC) (POTENTIAL).  
FT CONFLICT 599 599 R -> L (IN REF. 1).  
SC SEQUENCE 1134 AA: 1246+8 MW: 96126.63 kDa  
Quality Match 31.5%; Score 820.5; DB 1; Length 1134;  
Host local Similarity 36.0%; Pred. No. 146-57;  
Matches 169; Conservative 70; Mismatches 191; Indels 39; Gaps 7;  
QY 5 ASLVCGVSLLSGVGAMDLILLNSPLVSDAETSLTICIA---SGWRPHEPILIGRD 60  
DB 6 SSLLI--PTLPLASHVGASVDLTLLANLKTDPQRFHICVSGACAGRSSDPLILLK 63  
QY 61 -----FEALMNHQHPLEVTQVTPFWAKVKVWREKASKINAYPPRPVP 107  
DB 64 LGIVRTFPQQLYLAAPNASH-VILP-----GESKPSLLVGVESWVAG 108  
QY 108 GRAIPIPIKMKMGKASHFPAHLLIMIVKGVNINISPKKVLKEHIAVIYKNSFIHSVP 167  
DB 109 APRTVLYVHNSPGAHFPFKVTHTVNKGDTAVLSAHVHKRQTDVIMKNNSYNTLDW 168  
QY 168 HEVPO-ILEVHLPHPQJQJGVSAFYIGSLFTSAFTPLIVPPTEAKWGPENILCTA 226  
DB 169 QVADNPKPQQLQNVQPSLSIYSAIYFASPLQSAAPPILVPGGACAPWGCWVKIC 228  
QY 227 QNNNVWHELEPTTPFTMSTEFKVELHTTETLWESNGPFPFVYVLEFYG 286  
DB 229 TLFWWVWHPGQVYFTFTTETFEQA-PEGPGFSGVGEFETSTAYFELLLEFYG 288  
QY 287 QSCATGKMGKQNEACHPGYGYGCKLKCSNNHGMCKDFUSCLASPGWGLCKRGLP 346  
DB 289 QSGSGWQSGSLEA-APGHPJADPLQALNGSTGTFSGVTFSGWGVKESD-- 346  
QY 347 RMTPKIVDLPDHIENVSGKNEPI-CKASGWLPTNEETLVKPDGTVLHPKDNITDIHS 405  
DB 347 -KIQIISMAITVEFNIGIMPIVNCAAAQNPFPVQSMALPKPFGIMLSIKVIVPDKT 405  
QY 406 VALFTIHPILPDSGVWVGSVNTVAGWVEKPFENISVKVLPKPLNAPNVI 454  
DB 406 TAEFEVPSLTIHQSGPWFQVSTSGQSGPPRKPVKVVPPVPTTAPPELL 454  
RESULT 7  
SPEC\_HUMAN  
ID SPEC\_HUMAN STANDARD: PPT: 840 AA.  
AC Q14162; Q43701;  
DI 01 MAR 2002 (Ref. 41, Created)  
DI 01 MAR 2002 (Ref. 41, Last sequence update)  
DI 01 MAR 2002 (Ref. 41, Last annotation update)  
DE Endothelial cellis scavenger receptor protein (Atp-ty) LIL (non-prior).  
GN SPCB: KIAA0149.  
OS Homo sapiens (Human).  
OC Eukaryota, Metazoa, Chordata, Craniata, Vertebrata, Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE-Umbilical vein endothelial cells;  
RX MEDLINE=98058897; PubMed=9395444;  
RA Adachi H., Tsujimoto M., Arai H., Inoue K.;  
RT "Expression cloning of a novel scavenger receptor from human  
RT endothelial cells."  
PL J. Biol. Chem. 272:31217-31220(1997).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE-Bone marrow;





FT	DOMAIN	298	356	LAMININ EGF-LIKE 1.	FT	DISULFID	1461	1469	HY SIMILARITY.
FT	DOMAIN	357	426	LAMININ EGF-LIKE 2.	FT	DISULFID	1463	1476	HY SIMILARITY.
FT	DOMAIN	427	471	LAMININ EGF-LIKE 3.	FT	DISULFID	1478	1487	BY SIMILARITY.
FT	DOMAIN	472	518	LAMININ EGF-LIKE 4.	FT	DISULFID	1490	1503	BY SIMILARITY.
FT	DOMAIN	519	563	LAMININ EGF-LIKE 5.	FT	DISULFID	1506	1520	BY SIMILARITY.
FT	DOMAIN	564	609	LAMININ EGF-LIKE 6.	FT	DISULFID	1508	1527	BY SIMILARITY.
FT	DOMAIN	610	655	LAMININ EGF-LIKE 7.	FT	DISULFID	1529	1538	HY SIMILARITY.
FT	DOMAIN	656	700	LAMININ EGF-LIKE 8.	FT	DISULFID	1541	1551	HY SIMILARITY.
FT	DOMAIN	701	755	LAMININ EGF-LIKE 9.	FT	DISULFID	1554	1566	BY SIMILARITY.
FT	DOMAIN	756	808	LAMININ EGF-LIKE 10.	FT	DISULFID	1556	1573	BY SIMILARITY.
FT	DOMAIN	809	859	LAMININ EGF-LIKE 11. (INCOMPLETE).	FT	DISULFID	1575	1584	BY SIMILARITY.
FT	DOMAIN	1415	1460	LAMININ EGF-LIKE 12.	FT	DISULFID	1587	1602	HY SIMILARITY.
FT	DOMAIN	1461	1505	LAMININ EGF-LIKE 13.	FT	DISULFID	1630	1639	HY SIMILARITY.
FT	DOMAIN	1506	1553	LAMININ EGF-LIKE 14.	FT	DISULFID	1830	1839	HY SIMILARITY.
FT	DOMAIN	1554	1604	LAMININ EGF-LIKE 15.	FT	DISULFID	1832	1846	HY SIMILARITY.
FT	DOMAIN	1605	1614	LAMININ EGF-LIKE 16 (N-TERMINAL).	FT	DISULFID	1849	1858	BY SIMILARITY.
FT	DOMAIN	1615	1756	LAMININ DOMAIN IV.	FT	DISULFID	1861	1877	BY SIMILARITY.
FT	DOMAIN	1757	1829	LAMININ EGF-LIKE 16 (C-TERMINAL).	FT	DISULFID	1880	1894	BY SIMILARITY.
FT	DOMAIN	1830	1879	LAMININ EGF-LIKE 17.	FT	DISULFID	1882	1905	HY SIMILARITY.
FT	DOMAIN	1880	1936	LAMININ EGF-LIKE 18.	FT	DISULFID	1907	1916	HY SIMILARITY.
FT	DOMAIN	1937	1989	LAMININ EGF-LIKE 19.	FT	DISULFID	1919	1934	HY SIMILARITY.
FT	DOMAIN	1990	2046	LAMININ EGF-LIKE 20.	FT	DISULFID	1937	1951	BY SIMILARITY.
FT	DOMAIN	2047	2083	LAMININ EGF-LIKE 21.	FT	DISULFID	1939	1958	BY SIMILARITY.
FT	DOMAIN	2084	2121	LAMININ EGF-LIKE 22.	FT	DISULFID	1961	1970	BY SIMILARITY.
FT	DOMAIN	2693	2884	LAMININ G-LIKE 1.	FT	DISULFID	1973	1987	HY SIMILARITY.
FT	DOMAIN	2896	3056	LAMININ G-LIKE 2.	FT	DISULFID	1990	2000	HY SIMILARITY.
FT	DOMAIN	3072	3245	LAMININ G-LIKE 3.	FT	DISULFID	1992	2007	BY SIMILARITY.
FT	DOMAIN	3310	3482	LAMININ G-LIKE 4.	FT	DISULFID	2009	2018	BY SIMILARITY.
FT	DOMAIN	3498	3669	LAMININ G-LIKE 5.	FT	DISULFID	2021	2034	BY SIMILARITY.
FT	DISULFID	298	307	HY SIMILARITY.	FT	DISULFID	2037	2048	BY SIMILARITY.
FT	DISULFID	300	320	HY SIMILARITY.	FT	DISULFID	2039	2055	HY SIMILARITY.
FT	DISULFID	322	331	HY SIMILARITY.	FT	DISULFID	2057	2066	HY SIMILARITY.
FT	DISULFID	334	354	BY SIMILARITY.	FT	DISULFID	2069	2081	HY SIMILARITY.
FT	DISULFID	357	366	BY SIMILARITY.	FT	DISULFID	2084	2096	BY SIMILARITY.
FT	DISULFID	379	391	BY SIMILARITY.	FT	DISULFID	2086	2103	BY SIMILARITY.
FT	DISULFID	394	403	HY SIMILARITY.	FT	DISULFID	2105	2114	BY SIMILARITY.
FT	DISULFID	406	424	HY SIMILARITY.	FT	DISULFID	2117	2129	HY SIMILARITY.
FT	DISULFID	427	438	HY SIMILARITY.	FT	CARBOHYD	121	121	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	DISULFID	429	445	BY SIMILARITY.	FT	CARBOHYD	140	140	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	DISULFID	447	456	BY SIMILARITY.	FT	CARBOHYD	249	249	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	DISULFID	459	469	BY SIMILARITY.	FT	CARBOHYD	351	351	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	DISULFID	472	484	HY SIMILARITY.	FT	CARBOHYD	477	477	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	DISULFID	474	491	HY SIMILARITY.	FT	CARBOHYD	511	511	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	DISULFID	493	502	HY SIMILARITY.	FT	CARBOHYD	530	530	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	DISULFID	505	516	BY SIMILARITY.	FT	CARBOHYD	634	634	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	DISULFID	519	531	BY SIMILARITY.	FT	CARBOHYD	761	761	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	DISULFID	521	538	BY SIMILARITY.	FT	CARBOHYD	1014	1014	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	DISULFID	540	549	HY SIMILARITY.	FT	CARBOHYD	1341	1341	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	DISULFID	552	561	HY SIMILARITY.	FT	CARBOHYD	1705	1705	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	DISULFID	564	576	HY SIMILARITY.	FT	CARBOHYD	1756	1756	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	DISULFID	566	583	BY SIMILARITY.	FT	CARBOHYD	1868	1868	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	DISULFID	585	594	BY SIMILARITY.	FT	CARBOHYD	1944	1944	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	DISULFID	610	622	HY SIMILARITY.	FT	CARBOHYD	1986	1986	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	DISULFID	612	624	HY SIMILARITY.	FT	CARBOHYD	2002	2002	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	DISULFID	631	640	HY SIMILARITY.	FT	CARBOHYD	2159	2159	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	DISULFID	643	653	BY SIMILARITY.	FT	CARBOHYD	2207	2207	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	DISULFID	656	668	BY SIMILARITY.	FT	CARBOHYD	2231	2231	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	DISULFID	658	674	BY SIMILARITY.	FT	CARBOHYD	2235	2235	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	DISULFID	676	685	BY SIMILARITY.	FT	CARBOHYD	2401	2401	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	DISULFID	688	698	BY SIMILARITY.	FT	CARBOHYD	2421	2421	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	DISULFID	701	715	HY SIMILARITY.	FT	CARBOHYD	2487	2487	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	DISULFID	703	724	BY SIMILARITY.	FT	CARBOHYD	2821	2821	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	DISULFID	726	735	BY SIMILARITY.	FT	CARBOHYD	3087	3087	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	DISULFID	738	753	BY SIMILARITY.	FT	CARBOHYD	3242	3242	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	DISULFID	756	770	HY SIMILARITY.	FT	CARBOHYD	3541	3541	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	DISULFID	758	777	HY SIMILARITY.	SQ	SEQUENCE	3672	3672	AA: 404223 MW: 28E26,DB5FF14BFA CRC64:
FT	DISULFID	779	788	BY SIMILARITY.					
FT	DISULFID	791	806	BY SIMILARITY.					
FT	DISULFID	1415	1427	BY SIMILARITY.					
FT	DISULFID	1417	1444	BY SIMILARITY.					
FT	DISULFID	1436	1445	BY SIMILARITY.					
FT	DISULFID	1448	1458	BY SIMILARITY.					

Query Match 8.5% Score 221.0H 1: length 4672;

Best Local Similarity 24.9% Pred No. 3e-09;

Matches 68; Conservative 25; Mismatches 78; Indels 102; Gaps 14;

QY 193 YIGGNLFTSAPTRIVRRCCEAKWGPGCNH-----CIGAC-----MNN 240













```

FT DISULFID 479 488      BY SIMILARITY.
FT DISULFID 495 506      BY SIMILARITY.
FT DISULFID 500 515      BY SIMILARITY.
FT DISULFID 517 526      BY SIMILARITY.
FT DISULFID 533 588      BY SIMILARITY.
FT DISULFID 582 597      BY SIMILARITY.
FT DISULFID 599 608      BY SIMILARITY.
FT DISULFID 615 625      BY SIMILARITY.
FT DISULFID 619 634      BY SIMILARITY.
FT DISULFID 636 645      BY SIMILARITY.
FT DISULFID 652 663      BY SIMILARITY.
FT DISULFID 657 672      BY SIMILARITY.
FT DISULFID 674 683      BY SIMILARITY.
FT DISULFID 686 700      BY SIMILARITY.
FT DISULFID 695 709      BY SIMILARITY.
FT DISULFID 711 720      BY SIMILARITY.
FT DISULFID 803 814      BY SIMILARITY.
FT DISULFID 808 823      BY SIMILARITY.
FT DISULFID 825 834      BY SIMILARITY.
FT DISULFID 841 852      BY SIMILARITY.
FT DISULFID 846 865      BY SIMILARITY.
FT DISULFID 867 876      BY SIMILARITY.
FT DISULFID 883 894      BY SIMILARITY.
FT DISULFID 888 903      BY SIMILARITY.
FT DISULFID 905 914      BY SIMILARITY.
FT DISULFID 921 932      BY SIMILARITY.
FT DISULFID 926 941      BY SIMILARITY.
FT DISULFID 943 952      BY SIMILARITY.
FT CARBOHYD 152 152      N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 156 196      N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 247 247      N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 331 331      N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 412 412      N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 452 452      N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 558 558      N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 749 749      N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 965 965      N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 977 977      N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1034 1004      N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1040 1030      N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1150 1150      N-LINKED (GLCNAC. .) (POTENTIAL).
FT CONFLICT 14 17        MISSING (IN REF. 2).
FT CONFLICT 27 27        P -> A (IN REF. 2).
FT CONFLICT 1352 1352      T -> S (IN REF. 2).
SQ SEQUENCE 1408 AA: 15066) MW: 569044270997840 CPC264;

Query Match 7.7%; Score 199.5; DB 1; Length 1408;
Best Local Similarity 24.5%; Pred No 5 le-98;
Matches 76; Conservative 22; Mismatches 88; Indels 137; Gaps 17

QY 210 KQVAKWQPPCNH---LQI--ACMNQVCHHFGHICGICPGFMQKTCACEDLHPKTC 264
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 786 KIVGHITGECEINLNEISPNPCENGLTIDPFP-----IPE--GMSWTKRQ 834
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 265 KECGSG--QEGKSVYFPLP-----DPYGSATWKKGLQWNEA----- 301
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 835 SEKATGYAQGQNGTGMZAAKAAQIIPGAPWQGHICAPADQWQWQCHNDSI 893
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 402 -----CHPFPYIPYPI--PES-----NNSEM--DEFG CLSPFWGLGCE 341
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 894 GEGAGWRCVVAQPSRPPIINVNEISPAQCQATCTIGTGYCSITPPSPHGLPQF 953
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 342 -----PQIPMTIPK 352
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 954 ILISPKSAQNASNTISPTALNPSINMLDIALDHPIDEPENNAVQENNTSWNLW 1013
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 353 VILPCHIEVN---SGKENPITKASWPLPNEEMTLVKP-----DGLVLPKDFNHTDH 403
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1014 GGLPNYKVKVPSAKSNISGCVKQWQVVAISLSEISPNVKKCKALPS----- 1056
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 404 FSWAITTHRIPLPP---ESNVW 422
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

Db 1067 -----RRVAPPLPAKSSCW 1081

RESULT 14
ID WIF1_HUMAN STANDARD: PRT; 379 AA.
AC Q9Y5W5;
DI 16-OCT-2001 (rel. 40, created)
DI 16-OCT-2001 (rel. 40, last sequence update)
DI 16-OCT-2001 (rel. 40, last annotation update)
DE Wnt inhibitory factor 1 precursor (WIF-1).
GN WIF1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
SEQUENCE FROM N.A.
MEDLINE=99215557; PubMed=10201374;
RA Hsieh J.-C., Kodjabachian L., Rebertus M.L., Rattner A.,
RA Smallwood P.M., Samos C.H., Nusse R., Dawid I.H., Nathans J.;
RT "A new secreted protein that binds to Wnt proteins and inhibits their
RT activities."
RL Nature 398:431-436(1999).
CC -!- FUNCTION: BINDS TO WNT PROTEINS AND INHIBITS THEIR ACTIVITIES. MAY
CC BE INVOLVED IN MESODERM SEGMENTATION.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: CONTAINS 5 EGF-LIKE DOMAINS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed, usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/annouce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: AF122522; AAC25402.1; -
CC HSSP: P00743; LOCF.
CC MIM: 605186; -
CC InterPro: IPR000581; EGF-like.
CC InterPro: IPR002049; Laminin_EGF.
CC InterPro: IPR003306; WIF.
CC Pfam: PF00008; EGF; 5.
CC Pfam: PF02019; WIF; 1.
CC PRINTS: PR00011; EGF/LAMININ.
CC SMART: SM00469; WIF; 1.
CC PROSITE: PS00022; EGF_1; 5.
CC PROSITE: PS01186; EGF_2; 4.
KW Repeat; EGF like domain; Signal; Developmental protein.
FT SIGNAL 1 28 POTENTIAL.
FT CHAIN 29 379 WNT INHIBITORY FACTOR 1.
FT DOMAIN 177 208 EGF-LIKE 1.
FT DOMAIN 209 240 EGF-LIKE 2.
FT DOMAIN 241 272 EGF-LIKE 3.
FT DOMAIN 273 304 EGF-LIKE 4.
FT DOMAIN 305 336 EGF-LIKE 5.
FT DISULFID 177 186 POTENTIAL.
FT DISULFID 182 192 POTENTIAL.
FT DISULFID 198 200 POTENTIAL.
FT DISULFID 209 218 POTENTIAL.
FT DISULFID 214 224 POTENTIAL.
FT DISULFID 230 240 POTENTIAL.
FT DISULFID 241 250 POTENTIAL.
FT DISULFID 245 256 POTENTIAL.
FT DISULFID 262 264 POTENTIAL.
FT DISULFID 273 282 POTENTIAL.
FT DISULFID 278 288 POTENTIAL.
FT DISULFID 294 296 POTENTIAL.
FT DISULFID 305 314 POTENTIAL.
FT DISULFID 310 320 POTENTIAL.
FT DISULFID 326 328 POTENTIAL.

```

FT CAPB-HYD 88 N-LINKED (GLCNAC...) (POTENTIAL)  
 FT CAPB-HYD 245 N-LINKED (GLCNAC...) (POTENTIAL)  
 SQ SEQUENCE 379 AA: 41512 MW: 27782376264784 C6C4,  
 Query Match 7.6%; Score 198.5; DB:1; Length 379;  
 Best Local Similarity 26.4%; Pred No 1 ap-OR,  
 Matches 72; Conservative 23; Mismatches 95; Indels 81; Gaps 15,  
 QY 92 KASKINGAYCFGRVQ-PAIRIPIMKMQQASFLPATITMTVDKNDVNISFKKVLKE 150  
 DB 130 KASVQVQGFPCQKQGVAAFEVDVIMNSEGN-----TILQTP-----NALIFKICLOA 180  
 QY 151 EDVINYKNGEIIISVPHVEVDILEVILPHAPQDAGVGYIGNLTSAFTLIVRR 210  
 DB 181 PCPGCGNGGFGFNRHICGCPD--GFHGH----- 208  
 QY 211 CEAGKWKPGECNHLACI-ACMNGVCHEDTGEICGPRGEMGTCEKA-CELIIF-GRTC--K 265  
 DB 209 CE-----KALCTPPMNGVGLP-VLPGEIIPPGFYVNVTKAN-STI-FNGSTCTYP 259  
 QY 266 FRC-----SQDQDCKSYVCPDIP-----YQPSCATQWKGIQNP-ACHQPGVQDNC 311  
 DB 260 GKCTCPGTELEVELEISKPPKPNKAKGLGSKCKKCKSGYQDGLCKSPVCEFG----- 313  
 QY 312 KLPSGNNNGEWDPRFGQCLSPGWQGLQGER 342  
 DB 314 -----CGAHGTCHRPKNCQCGQGHGHCNK 339  
 RESULT 15  
 DL\_DROME STANDARD: PRT: 833 AA.  
 ID P10041; Q9V0Y2; Q99108;  
 AC 01-MAR-1989 (Rel. 10, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DE 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Neurogenic locus Delta protein precursor.  
 GN DL OR C33619.  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota, Metazoa, Arthropoda, Tracheata, Hexapoda, Insecta,  
 OC Pterygota, Neoptera, Endopterygota, Diptera, Brachycera, Muscophora,  
 OC Ephydroidea, Drosophilidae; Drosophila  
 OX NCBI\_TaxID=7227;  
 RN [1]  
 RP SEQUENCE FROM N.A., TISSUE SPECIFICITY, AND DEVELOPMENTAL STAGE.  
 RC TISSUE=Embryo;  
 RA Vaessin H., Bremer K.A., Knust E., Campos-Ortega J.A.;  
 RT "The neurogenic gene Delta of Drosophila melanogaster is expressed in  
 neurogenic territories and encodes a putative transmembrane protein  
 with EGF-like repeats.";  
 RI EMBO J. 6:3431-3440(1987).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC SPAIN-OREGAN P. TISSUE=Embryo,  
 RX MEDLINE=89196890; PubMed=1149249;  
 RA Kopynski C.C., Altan K., Rechter K., Koch P.J., Maskavitch M.A.F.,  
 RI "Delta, a Drosophila neurogenic gene, is transcriptionally complex and  
 encodes a protein related to blood coagulation factors and epidermal  
 growth factor of vertebrates.";  
 RI Genes Dev. 2:1723-1735(1988).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BERKELEY;  
 RX MEDLINE=20196096; PubMed=10731132;  
 RA Adams M.D., Gelinkin S.E., Hill F.A., Evans C.A., Gwynne T.D.,  
 RA Ananides P.G., Scherer S.E., Li P.W., Hoskins P.A., Galle P.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.G., Wortman J.P., Vandell M.D., Zhang Q., Chen L.X.,  
 RA Brandon P.C., Rogers Y.-H.C., Blarck P.C., Chang M., Pfeiffer H.P.,  
 RA Wan K.H., Doyle C., Hyatt P.C., Heit G., Nelson C.K., Miklos G.L.G.,  
 RA April J.F., Aghayani A., An H.-J., Andrews-Pfankuch C., Baldwin B.,  
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,

Reeson K.Y., Benes P.V., Berman B.P., Bhandari D., Boelschewsky S.,  
 Burkova D., Butcher M.P., Bouck J., Brokstein P., Brothier P.,  
 Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 Cherry J.M., Cawley S., Dahlke C., Davenport L.H., Davies P.,  
 de Paolis R., Delcher A., Deng Z., Mays A.D., New J., Dietz S.M.,  
 Dodson K., Doup L.E., Downes M., Duquenois S., Dunkov B.C., Dunn P.,  
 Durbin K.J., Evangelista C.C., Ferrar C., Ferreira S., Fleischmann W.,  
 Fuster G., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
 Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 Harris N.J., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
 Hostin G., Houston K.A., Howland T.J., Wei M.H., Isonwam C.,  
 Kalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 Liu X., Mattei H., McIntosh T.C., McLeod M.P., McPherson D.,  
 Mckelvey G., Mijishina N.V., Mobarry C., Morris J., Mushrefi A.,  
 Mount S.M., Moy M., Murphy R., Murthy L., Muzny D.M., Nelson D.L.,  
 Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paolel J.M.,  
 Palazzolo M., Pittman G.S., Fan S., Pollard J., Puri V., Reese M.G.,  
 Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 Spier P., Spradling A.C., Stapleton M., Strong P., Sun E.,  
 Svirska K., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 Ye J., Yeh P.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 Zheng X.H., Zhou F.N., Zhou W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RT "The genome sequence of Drosophila melanogaster.";  
 RI Science 287:2185-2195(2009).  
 PI [4]  
 RP SEQUENCE OF 422 621 FROM N.A.  
 RC TISSUE=Embryo;  
 RX MEDLINE=87218537; PubMed=3107986;  
 RA Knust E., Dietrich U., Tepass U., Bremer K.A., Weigel D., Vaessin H.,  
 Campos-Ortega J.A.;  
 RT "ECF homologous sequences encoded in the genome of Drosophila  
 melanogaster, and their relation to neurogenic genes.";  
 RI EMBO J. 6:761-766(1987).  
 RN [5]  
 RP PATTERN OF TRANSCRIPTION, AND CHARACTERIZATION.  
 RX MEDLINE=91299246; PubMed=2128477;  
 RA Haenlin M., Gramateschek B., Campos-Ortega J.A.;  
 RT "The pattern of transcription of the neurogenic gene Delta of  
 Drosophila melanogaster.";  
 RI Development 110:905-914(1990).  
 CC -1- FUNCTION: ESSENTIAL FOR PROPER DIFFERENTIATION OF ECTODERM. DE  
 CC DEVELOPMENT IS REQUIRED FOR THE CORRECT SEPARATION OF NEURAL AND EPIDERMAL  
 CC CELL LINEAGES.  
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC ABILITIES, FOR EXAMPLE THE NEUROGENIC ECTODERM AND THE PRIMITIVE  
 CC OF THE SENSE ORGANS. LATER EXPRESSION IS RESTRICTED TO THOSE CELLS  
 CC THAT HAVE ADOPTED A NEURAL FATE.  
 CC -1- DEVELOPMENTAL STAGE: EXPRESSED BOTH MATERNALLY AND ZYGOTICALLY.  
 CC EXPRESSION IS HIGHEST EARLY IN EMBRYONIC DEVELOPMENT (STAGE 5) AND  
 CC REDUCES TO A LOW LEVEL DURING LARVAL STAGES.  
 CC -1- MISCELLANEOUS: SEPARATION OF NEUROBLASTS FROM THE ECTODERM INTO  
 CC THE INNER PART OF EMBRYO IS ONE OF THE FIRST STEPS OF CNS  
 CC DEVELOPMENT IN INSECTS, THIS PROCESS IS UNDER CONTROL OF THE  
 CC NEUROGENIC GENES.  
 CC -1- MISCELLANEOUS: NOTCH AND SERRATE MAY INTERACT AT THE PROTEIN  
 CC LEVEL, IT IS CONSIDERABLE THAT THE SERRATE AND DELTA PROTEINS MAY  
 CC COMPLETE FOR BINDING WITH THE NOTCH PROTEIN.  
 CC -1- SIMILARITY: CONTAINS 9 EGF-LIKE DOMAINS.  
 CC -1- SIMILARITY: BELONGS TO THE DELTA/SERRATE/NOTCH FAMILY.  
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb.scrib.org/announce/>)







KW Tyrosine-protein kinase.  
 SQ SEQUENCE 1083 AA: 12202P MW: 649750 Da P4430 CR644

Query Match 86.7%; Score 2258; DB 11; Length 1083.  
 Best Local Similarity 93.18; Pred. No. 2,70-211;  
 Matches 403; Conservative 12; Mismatches 18; Indels 0; Gaps 0;

QY 40 TSLTCLASWPEPITIGPEALMNAJLQGP[FVTLGVPEWAKKVVVKKKASKINGA 99  
 DQ 1 TSLTCLASWPIETITIGPEALMNAJLQGP[FVTLGVPEWAKKVVVKKKASKINGA 60  
 QY 100 YFEGVGRGPAIRIRIMMKKQASHLPATLIMIVUKGHNVINISKKVLKKEEDAVLYKNG 159  
 DQ 61 YFEGVGRGPAIRIRIMMKKQASHLPATLIMIVUKGHNVINISKKVLKKEEDAVLYKNG 120  
 QY 160 SFTHSVPRHEVPELLEVLPHAPUPJUNAVYSARYIGGSLFTSAPTRILVPRCEAKWKWPE 219  
 DQ 121 SFTHSVPRHEVPELLEVLPHAPUPJUNAVYSARYIGGSLFTSAPTRILVPRCEAKWKWPE 180  
 QY 220 CNHLCTAEMNNVWVHEFJAEVTPPEFMSPTEFANTELHTEPPTVEPSSQEPKSYVP 279  
 DQ 181 CNRPCTTKNNVWVHEFJAEVTPPEFMSPTEFANTELHTEPPTVEPSSQEPKSYVP 240  
 QY 280 CLDHPKATSTAGWKGQVNFANFAPGPGSGPKGKQKQSTNNPMMVUTPGTGLSPGWAQGLQ 339  
 DQ 241 CLDHPKATSTAGWKGQVNFANFAPGPGSGPKGKQKQSTNNPMMVUTPGTGLSPGWAQGLQ 300  
 QY 340 CEREGLPRMTPTKIVLPDPIHTEVNSKFNPTKASGWPLTNEEMTLVKPGDTVLBPKNDFN 399  
 DQ 301 CKEGHPMTPTKIVLPDPIHTEVNSKFNPTKASGWPLTNEEMTLVKPGDTVLBPKNDFN 360  
 QY 400 HTDFESVAIFTHIRLPDPSGWWGCVNTVAGWVEKPNISVKVLPKPLNAPNYDITGHN 459  
 DQ 461 HTDFESVAIFTHIRLPDPSGWWGCVNTVAGWVEKPNISVKVLPKPLNAPNYDITGHN 420  
 QY 460 FAVINISSEPYEG 472  
 DQ 421 FAVINISSEPYEG 433

RESULT 2  
 073791 PRELIMINARY; PRI: 1116 AA.

AC 073791  
 DT 01-AUG-1998 (Tremblrel. 07, Created)  
 DT 01-AUG-1998 (Tremblrel. 07, Last sequence update)  
 DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)  
 DE ANGIOPOIETIN 1 RECEPTOR PRECURSOR (PC 2.7.1.112) (TYROSINE-PROTEIN KINASE RECEPTOR FIE-2) (ENDOTHELIAL-SPECIFIC RECEPTOR TYROSINE KINASE TIE-2)  
 GN TIE2 (p TIE2 (p TIE-2)  
 OS Brachydanio rerio (Zebrafish) (Zebra danio).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi; Cypriniformes; Cyprinidae; Danio.  
 OX NCBI\_TaxID:7955;  
 RN [1]  
 RP SEQUENCE FROM N.A. AND TISSUE SPECIFICITY  
 RC TISSUE:EMBRYO;  
 RA MEDLINE:98264593; PubMed:9603430;  
 RX Lyons M.S., Bell R., Stainer D., Peters K.G.;  
 RT Isolation of the zebrafish homologues for the tie-1 and tie-2 endothelium-specific receptor tyrosine kinases.;  
 RL Dev. Dyn. 212:133-140(1998).

CC -!- FUNCTION: THIS PROTEIN IS A PROTEIN TYROSINE KINASE. TRANSMEMBRANE RECEPTOR FOR ANGIOPOIETIN 1. PROMOTELY PROLIFERATES ENDOTHELIAL CELL PROLIFERATION, DIFFERENTIATION AND GUIDES THE PROPER PATTERNING OF ENDOTHELIAL CELLS DURING BLOOD VESSEL FORMATION (BY SIMILARITY).  
 CC -!- CATALYTIC ACTIVITY: ATP + A PROTEIN TYROSINE - ADP + PROTEIN TYROSINE PHOSPHATE.  
 CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.  
 CC -!- TISSUE SPECIFICITY: EXPRESSED IN MOST POPULATIONS OF ENDOTHELIAL.

CC 1 CELLS IN 24 HOUR EMBRYOS. NOT PRESENT IN INTERSEGMENTAL VESSELS.  
 CC 1 SIMILARITY TO OTHER PROTEIN TYROSINE KINASES IN THE CATALYTIC DOMAIN.  
 CC -!- SIMILARITY: CONTAINS 2 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.  
 CC -!- SIMILARITY: CONTAINS 3 FIBRONECTIN TYPE III-LIKE DOMAINS.  
 DR EMBL: AF053632; AAC04331 1; 2;  
 DR HSSD: P11362; 1PKC;  
 DR 2FIN: Z59-GENE-990415-56; tie2;  
 DR InterPro: IPR000561; EGF-like;  
 DR InterPro: IPR000719; EGF\_kinase;  
 DR InterPro: IPR003961; FN\_III;  
 DR InterPro: IPR003599; Ig;  
 DR InterPro: IPR001245; Tyr\_kinase;  
 DR Pfam: PF00041; t03; 3;  
 DR Pfam: PF00069; kinase; 1;  
 DR PRINTS: PR00109; TYRKINASE;  
 DR SMART: SM00181; EGF; 2;  
 DR SMART: SM00050; FN3; 2;  
 DR SMART: SM00409; IG; 1;  
 DR SMART: SM00219; TyKc; 1;  
 DR PROSITE: PS00022; EGF\_1; DNKNWR\_2;  
 DR PROSITE: PS01186; EGF\_2; 2;  
 DR PROSITE: PS00107; PROTEIN\_KINASE\_ATP; 1;  
 DR PROSITE: PS00011; PROTEIN\_KINASE\_DOM; 1;  
 DR PROSITE: PS00109; PROTEIN\_KINASE\_TYR; 1;  
 KW Receptor; Tyrosine-protein kinase; Transmembrane; Signal; ATP-binding;  
 KW Repeat; EGF-like domain; Transmembrane; Immunoglobulin domain;  
 KW Glycoprotein; Phosphorylation; Multigene family;  
 FT SIGNAL 1 21 POTENTIAL;  
 FT CHAIN 22 1116 ANGIOPOIETIN 1 RECEPTOR;  
 FT DOMAIN 22 745 EXTRACELLULAR (POTENTIAL);  
 FT TRANSMEM 767 766 POTENTIAL;  
 FT DOMAIN 767 1116 CYTOPLASMIC (POTENTIAL);  
 FT DOMAIN 45 105 IG-LIKE C2-TYPE DOMAIN;  
 FT DOMAIN 214 256 EGF-LIKE 1;  
 FT DOMAIN 258 302 EGF-LIKE 2;  
 FT DOMAIN 304 342 EGF-LIKE 3;  
 FT DOMAIN 368 422 IG-LIKE C2-TYPE DOMAIN;  
 FT DOMAIN 442 535 FIBRONECTIN TYPE-III;  
 FT DOMAIN 538 627 FIBRONECTIN TYPE-III;  
 FT DOMAIN 631 727 FIBRONECTIN TYPE-III;  
 FT DOMAIN 816 1088 PROTEIN KINASE;  
 FT NP\_BIND 822 830 ATP (BY SIMILARITY);  
 FT BINDING 847 847 ATP (BY SIMILARITY);  
 FT ACT\_SITE 956 956 BY SIMILARITY;  
 FT MOD\_RES 984 984 PHOSPHORYLATION (AUTO-) (BY SIMILARITY);  
 FT MOD\_RES 1094 1094 PHOSPHORYLATION (AUTO-)  
 FT MOD\_RES 1105 1105 PHOSPHORYLATION (AUTO-)  
 FT CARBOHYD 110 110 N-LINKED (GLCNAC...)  
 FT CARBOHYD 143 143 N-LINKED (GLCNAC...)  
 FT CARB-HYD 223 223 N-LINKED (GLCNAC...)  
 FT CARBOHYD 367 367 N-LINKED (GLCNAC...)  
 FT CARBOHYD 387 387 N-LINKED (GLCNAC...)  
 FT CARBOHYD 425 425 N-LINKED (GLCNAC...)  
 FT CARBOHYD 590 590 N-LINKED (GLCNAC...)  
 FT CARBOHYD 637 637 N-LINKED (GLCNAC...)  
 FT CARBOHYD 642 642 N-LINKED (GLCNAC...)  
 SQ SEQUENCE 1116 AA: 122360 MW: 441480 Da P4430 CR644

Query Match 34.5%; Score 899.5; DB 13; Length 1116;  
 Best Local Similarity 38.38; Pred. No. 1,3e-78;  
 Matches 188; Conservative 88; Mismatches 172; Indels 43; Gaps 15;

QY 1 MDSIASIVAGVSVLLSGTVEGAMDLIIINSIPVSD-AETSLRCLASGW----- 49  
 DQ 4 LDSQIALLLIG--CWSSGSAVLSIVHVN-PPVVSPTAFSLICVSSDSSGSGSVLAIG 61  
 QY 50 ----RPH-EPITIGDFFALMNQHDPLEVTDVTREKAKVWKREKASKINAYCEG 104  
 DQ 62 QEFPRPGSVIALGQEP-----PHEPRPHPAATVTVSSR-----SHAFGAPYQI 108

QY 105 PVPEAEIPITMKMPQASLPAILMTVTKVDNWNISFKKVLKEEDAVIYKNGFIHS 164  
 DB 109 R-INSTGKVIYKMIQAAFLPESLIIIVNQGNINISYRRILYSPEVIVTKNGHPS 167  
 QY 165 VPHFVPHIIEVHLPHAPQD-ACVYSARYICNIETISAFTRIIVRCAQKWCPCNHL 223  
 DB 168 SPEDISDITHYPTVTKAESHGTAIPIYISNAPSSAATPLIVISVARGWGFNTES 227  
 QY 224 CTACMNCVCHENVQITQVYQMCPTCKACPLIFGRTCKPCQSQGQ-CKSYVFCIP 282  
 DB 228 CPKCAAGCVNDITGCVCTGFRGHCIDIVCGHCFGACCKPC -VGVGCPALVFCIK 285  
 QY 283 LPVPS-AT-CK-GLY-NEA-HP-PP-GLP-KLPS-NN-EM-HP-PP-L-SP-WHL-CTER 442  
 DB 286 DPVGS-SCASWPCISNDACDNYGAGGCTKCKVAKGR-CHFFPSCV-A-CHGSSMDE 443  
 QY 343 EGIPMTKIVGLGDRHIVNSG KFNICKASGMPLINEMILVKKPDGIVLHPKFNHI 401  
 DB 344 AD---SSPVISHLRD-VEINTGVELSVNCSAGPPAPLHGTITLITANTTAAVDHTL 409  
 QY 402 DHSVAITTHRIIPHSVWVSVNIVAGMVKPKNISVAVLKPPIKINAFNVIDIGHNFA 461  
 DB 400 NUGSTSVFVQVQVSSAGRWKVCVNNTHMVQVDEIVLAKVVPKPCNPVLQSGCHRIH 459  
 QY 462 VINISSEPPYR 472  
 DB 460 LLIINTPEYSG 470

RESULT 4  
 Q9TQV2  
 ID Q9TQV2 PRELIMINARY: PRT: 434 AA.  
 AC Q9TQV2  
 DT 01-MAY-2000 (11PMHirel 13, Created)  
 DT 01-DEC-2001 (11EMHirel 13, Last sequence update)  
 DE CG18172 PROTEIN.  
 GN CG18172  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Insecta; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 OX NCBI\_TaxID=7227;  
 RN [1]  
 RP  
 RC STRAIN=BERKELEY;  
 PX MEDLINE=20196006; PubMed=107311132;  
 RA Adams M.D., Celisner S.E., Holt P.A., Evans C.A., Gocayne J.D.,  
 RA Ananides P.G., Scherer S.E., Li P.W., Hoskins P.A., Gallo P.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.G., Wurtman J.P., Vandeil M.B., Zhang Q., Chen J.X.,  
 RA Randon P.C., Rogers Y.-H., Blasej K.G., Champ M., Weisner R.D.,  
 RA Wan K.H., Boyle C., Baxter E.G., Holt G., Nelson C.K., Miklos G.L.G.,  
 RA Abrell J.F., Aghayani A., Au H.J., Andrews-Plamich R., Baldwin D.,  
 RA Balow R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Beeson K.Y., Benos P.V., Bergan P.P., Bhandari D., Bolshakov S.,  
 RA Borkova D., Borhan M.P., Brock J., Brockstein P., Brothman P.,  
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 RA Cherry J.M., Ciewley S., Dahlke C., Davernport L.B., Davies P.,  
 RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Doup L.E., Downes M., Dugan-Pocha S., Dunkov B.C., Dunn P.,  
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
 RA Foster C., Gabrielson A.E., Garg N.S., Gelbart W.M., Glasser K.,  
 RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.L., Harvey D., Herman T.J., Hernandez J.R., Houck J.,  
 RA Hostin D., Houston K.A., Howland T.L., Wei M.-H., Ibegwam C.,  
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mattei R., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Morkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
 RA Mount S.M., Mny M., Murphy L., Murphy L.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Parleth J.M.,

PA Palacchio M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Pelinert K., Remington K., Saunders P.C., Scheeler F., Shen H.,  
 RA Shue B.C., Stien-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun H.,  
 RA Svirskaas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissbach T.,  
 RA Williams S.M., Woodage T., Worley K., Wu C., Yang S., Yan Q.A.,  
 RA Ye J., Yeh P.-F., Zaveri T.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zhang X.H., Zheng F.N., Zheng W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 FT "The genome sequence of Drosophila melanogaster";  
 PL Science 287:2185-2195(2000).  
 DB EMBL: AE003472; ANAF5531; --  
 DB HSSP: P02468; IKLO.  
 DE FlyBase: FBgn035261; 2518172.  
 DR InterPro: IPR001774; DSL.  
 DR InterPro: IPR000561; EGF-like.  
 DR InterPro: IPR002049; Laminin\_EGF.  
 DR Pfam: PF00008; EGF; 8.  
 DR PRINTS: PR00011; EGF\_LAMININ.  
 DR SMART: SM00051; DSL; 1.  
 DR SMART: SM00180; EGF\_Lam; 3.  
 DR SMART: SM00001; EGF\_Like; 6.  
 DR PROSITE: PS00022; EGF\_1; 7.  
 DR PROSITE: PS01186; EGF\_2; 9.  
 DR EGF-like domain; Glycoprotein.  
 KW EGF-like domain; Glycoprotein.  
 SQ SEQUENCE 434 AA; 46417 MW; 26H5D1P5C7470E4 CRC64;

Query Match 12.6%; Score 327.5; DB 5; Length 434;  
 Best Local Similarity 35.8%; Pred No. 26e-24;  
 Matches 62; Conservative 26; Mismatches 74; Indels 11; Gaps 5;

QY 209 PPFPAKWDPCENHUTACM-NNWV-HEHDDPCVTHGEMCHDCKKAPHPGPCTCKR 267  
 DB 204 PFCANKY-LLCNGPT-E'DMEHTLD-HPPTFN-NT-LSKSNAG-TTPPTFLPYVFNDEL 463  
 QY 268 CSQEGCKSYVFCIPDFYCKSCATCKGLQNEAHPFYF-TPCKLRSCNSENEMCTDFQ 327  
 DB 264 CN-----CKMAKAKSVNGLICAPMPGHCESCEPSGFGACALRDCUNGAKCRPET 319  
 QY 328 G-CL-SPGQWGLKEP----P-IPMTKIVLHPKFNISGKFNICKASG 375  
 DB 320 GGLCTAGSKNKKICDRPCDNLNHFQGCCKVLCVCHNNAACNPUNGSLTC-AAG 371

RESULT 4  
 Q9TQV2  
 ID Q9TQV2 PRELIMINARY: PRT: 1654 AA.  
 AC Q9TQV2  
 DT 01-MAY-2000 (11PMHirel 13, Created)  
 DT 01-DEC-2001 (11EMHirel 13, Last sequence update)  
 DE Y64310A 7 PPT-TEIN.  
 GN Y64310A 7  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidae;  
 OC Rhabditidae; Rhabditinae; Caenorhabditis.  
 OX NCBI\_TaxID=6239;  
 RN [1]  
 RP  
 RC SEQUENCE FROM N.A.  
 PX MEDLINE=99069613; PubMed=9851916;  
 RA Mortimore B.J.;  
 RA Mortimore B.J.;  
 RA Submitted (APR-1999) to the EMBL/GenBank/DDJ databases.  
 RN [2]  
 RP  
 RC SEQUENCE FROM N.A.  
 PX MEDLINE=99069613; PubMed=9851916;  
 RA none;  
 FT "Genome sequence of the nematode C. elegans: A platform for  
 RT investigating biology.";  
 KL Science 282:2012-2018(1998).  
 RN [3]  
 RP  
 RC SEQUENCE FROM N.A.  
 PX Ainscough P.;

Submitted (MAY 1999) to the EMBL/GenBank/Joint databases  
 DR EMBL: AL117206; CAB60454.1; -  
 DR EMBL: AL110498; CAB60454.1; JOINED.  
 DR EMBL: AL110498; CAB57911.1; -  
 DR EMBL: AL117206; CAB57911.1; JOINED.  
 DR HSPD: P00736; IAPQ.  
 DR InterPro: IPR000152; ASX\_Hydroxyl.  
 DR InterPro: IPR000561; EGF-like  
 DR InterPro: IPR001881; EGF\_CA  
 DR InterPro: IPR002049; Laminin\_EGF.  
 DR Pfam: PF00008; EGF; 25.  
 DR PRINTS: PR00011; EGF\_LAMININ.  
 DR SMART: SM00179; EGF\_CA; 4.  
 DR SMART: SM00001; EGF\_Like; 18.  
 DR PROSITE: PS00010; ASX\_HYDROXYL; 4.  
 DR PROSITE: PS00023; EGF\_1; "NKR-WN\_22".  
 DR PROSITE: PS01186; EGF\_2; 24.  
 DR PROSITE: PS01187; EGF\_CA; 3.  
 DR Calcium-binding: EGF-like domain; Glycoprotein; Hydroxylation; Repeat.  
 KW SEQUENCE 1664 AA; 179279 MW; A6609384C705832 CRC64;

Query Match 12.38; Score 321; DB 5; Length 1664;  
 Best Local Similarity 34.48; Pred. No. 7.1e-22;  
 Matches 63; Conservative 22; Mismatches 66; Indels 32; Gaps 7.  
 QY 211 CEAAQKWGPCNHLCTACMNGVCHEDTQFCICPPGPGMGTCEKACELFTFGRTCKKRCSG 270  
 DB 896 CHGPGGSHWQDQKUNALQISNLSFAPQWSKKQDRAAPITPKKSKKQD 947  
 QY 271 QGQKSYVFLPLFVNGSNTATGKGLQNEAHFSTVTEGKLEKLSNNAE MLHLQ 328  
 DB 948 ---CAQRMHLPQSGEITPPKPKHKDETUSHLFSAFKGDSQUNATDSVTGSQ 1004  
 QY 330 LCPSPQWQGLQCHERIGIPRMTPKIVDLPHLIEVNSGKFNPLCK 371  
 DB 1005 PCFPCWPKKKCKLP-----PCPH-----GPFGRGCAICDCT-----TINDTSM 1042  
 QY 386 VKP 388  
 DB 1043 YNP 1045

RESULT 5  
 ID 088281 PRELIMINARY; PRT; 1574 AA.  
 AC 088281;  
 DT 01-NOV-1998 (TRIMBLrel. 08, Created)  
 DT 01-NOV-1998 (TRIMBLrel. 08, Last sequence update)  
 DT 01-DEC-2001 (TRIMBLrel. 19, Last annotation update)  
 DE MEGF6.  
 GN MEGF6.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC MEDLINE=98360089; PubMed=9693030;  
 RA Nakayama M., Nakajima D., Nakajima D., Kikuno K., Ohara O.  
 RT "Identification of high-molecular-weight proteins with multiple EGF-like motifs by motif-trap screening".  
 RL Genomics 51:27-34 (1998).  
 DR EMBL: AB011552; BAA32452.1; -  
 DR HSPD: P00736; IAPQ.  
 DR InterPro: IPR000152; ASX\_Hydroxyl.  
 DR InterPro: IPR000561; EGF-like.  
 DR InterPro: IPR001881; EGF\_CA  
 DR Pfam: PF00008; EGF; 24.  
 DR SMART: SM00179; EGF\_CA; 4.  
 DR SMART: SM00001; EGF\_Like; 19.  
 DR PROSITE: PS00010; ASX\_HYDROXYL; 5.

DR PROSITE: PS00023; EGF\_1; UNKNOWN\_23.  
 DR PROSITE: PS01186; EGF\_2; 23.  
 DR PROSITE: PS01187; EGF\_CA; 5.  
 KW Calcium-binding: EGF-like domain; Glycoprotein; Hydroxylation; Repeat.  
 KW SEQUENCE 1574 AA; 165445 MW; 2B48533DBF77F6E7 CRC64;  
 Query Match 12.28; Score 317.5; DB 11; Length 1574;  
 Best Local Similarity 35.08; Pred. No. 1.4e-21;  
 Matches 57; Conservative 21; Mismatches 62; Indels 24; Gaps 4.  
 QY 211 CEAAQKWGPCNHLCTACMNGVCHEDTQFCICPPGPGMGTCEKACELFTFGRTCKKRCSG 270  
 DB 777 CPEGWGHGCGHICPACHGASGCPETCTCLCLPFCVCSRCQDICSAGWYCTGCTIRCA 845  
 QY 271 QGQKSYVFLPLFVNGSNTATGKGLQNEAHFSTVTEGKLEKLSNNAE MLHLQ 328  
 DB 846 ---FANDGHCPTTGTGSCAPGWTCLSCQPAQHSCHWGPDPCTHMCNSAGHNGDAVSC 892  
 QY 329 CLCSPQWQGLQCHERIGIPRMTPKIVDLPHLIEVNSGKFNPLCK 371  
 DB 893 CLCEAYESPRKEQ-----SCQGYGFSCE 918

RESULT 6  
 ID 096K67 PRELIMINARY; PRT; 1140 AA.  
 AC 096K67;  
 DT 01-DEC-2001 (TRIMBLrel. 19, Created)  
 DT 01-DEC-2001 (TRIMBLrel. 19, Last sequence update)  
 DT 01-DEC-2001 (TRIMBLrel. 19, Last annotation update)  
 DE MEGF10 PROTEIN (KIAA1780).  
 GN MEGF10.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=HIPPOCAMPUS;  
 RC MEDLINE=21245130; PubMed=11347906;  
 RA Nagase T., Nakayama M., Nakajima D., Kikuno K., Ohara O.  
 RT "Prediction of the coding sequences of unidentified human genes. XX. The complete sequences of 100 new cDNA clones from brain which code for large proteins in vitro".  
 RL DNA Res. 8:85-95(2001).  
 DR EMBL: AB058676; BAB47404.1; -  
 DR SEQUENCE 1140 AA; 122204 MW; 45M2FA219412895A CRC64;

Query Match 12.18; Score 316; DB 4; Length 1140;  
 Best Local Similarity 29.48; Pred. No. 1.3e-21;  
 Matches 67; Conservative 13; Mismatches 58; Indels 90; Gaps 6.  
 QY 159 FTSATRIIVKKAQAGWGPENHLCTACMNGVCHEDTQFCICPPGPGMGTCEKACELFTFGRTCKKRCSG 271  
 DB 216 YTGAFCEDL---CPGKIGSPQCEQRCPQNGVGVCHIVIGECSPSGWMTVAGQPCPEG 271  
 QY 244 -----PFGMGTCEKAELITGTCTKEFGSGEGSK 275  
 DB 272 PEGKNTSQVQTHNQTAAAGGEGSYVQVQVIVGLVGVFAELQVNSQK 331  
 QY 276 SY-----VFCLDPFG-----CSQATGW 294  
 DB 332 CYHVSACLCEAGFAGECEARLCEGLYGIKCDKRCPEHLENTSHIPMSGECAKPGW 391  
 QY 284 KGIQCNFACHGPFYDPAKLPSCSNGNFMITLRFUGCLGSPQWQGLQ 340  
 DB 302 SGIYNETSTFTEACACGQICSTNADQLSVTSGTATAGFGAGLQ 439  
 RESULT 7  
 QY6K6b

ID Q96K06 PRELIMINARY: PRT: 969 AA.  
 AC Q96K06;  
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
 DE MGFI1 PROTEIN (K1AA1781).  
 GN MGFI1.  
 OS Homo sapiens (Human).  
 OC Eukaryota, Metazoa, Chordata, Craniata, Vertebrata, Euteleostomi,  
 OC Mammalia, Eutheria, Primates, Catarrhini, Hominoidea, Homo.  
 OX NCBI\_TaxID=9606  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=BRAIN;  
 RX MEDLINE=21245130, PubMed 11447406;  
 RA Nagase T., Nakayama M., Nakajima D., Kikuno R., Ohara O.;  
 RT "Prediction of the coding sequences of unidentified human genes. XX.  
 RT the complete sequences of 106 new cDNA clones from brain which code  
 RT for large proteins in vitro";  
 RL DNA Res. 8:85-95(2001).  
 DR EMBL: AB058677; BAB47410.1; .  
 SQ SEQUENCE 969 AA, 10165 MW, 5560.4 PFEI GCR209 G664.

Query Match 12.1%, Score 314, DB 4, Length 969,  
 Best Local Similarity 41.6%, Pred. No. 1.6e-21;  
 Matches 57, Conservative 14, Mismatches 52, Indels 14, Gaps 5;

QY 211 CEAKKWPENHLCTAMNNVWVHEDTSEICPPPMPTREKACELHIFGPTKEP--C 268  
 DB 405 CPGTGLNCSCT-CANCAACSPIDGSCCTPCWIGDTCCLPCDGFGLNCSHCDC 463  
 QY 269 SQBGGKSYVFLPDY--GGSATGWKGLQNEACHPGFGYGDCKLPKSCNNGEMCDRF 326  
 DB 464 SHADGC-----DPVTGTCCLAGTGTGTDSTCPGPGWPGNSVSCSNGSGSCSPE 515  
 QY 327 QC-CICSPGMOGLOQR 342  
 DB 516 DGSCEAFGFGFLQR 532

RESULT 8  
 Q14162  
 ID Q14162 PRELIMINARY: PRT: 830 AA.  
 AC Q14162;  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
 DE KIAA0149 PROTEIN.  
 GN KIAA0149.  
 OS Homo sapiens (Human).  
 OC Eukaryota, Metazoa, Chordata, Craniata, Vertebrata, Euteleostomi,  
 OC Mammalia, Eutheria, Primates, Catarrhini, Hominoidea, Homo.  
 OX NCBI\_TaxID=9606,  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC MEDLINE=96127530, PubMed-8590280;  
 RA Nagase T., Seki N., Tanaka A., Ishikawa K., Nomura N.;  
 RT "Prediction of the coding sequences of unidentified human genes. IV.  
 RT the coding sequences of 40 new genes (K1AA0121-K1AA0160) deduced by  
 RT analysis of cDNA clones from human cell line KG-1";  
 RL DNA Res. 2:167-174(1995)  
 DR EMBL: D63483; BAA09770.1; .  
 DR HSSP: F01140, 2BN2.  
 DR InterPro: IPR000561; EGF-like.  
 DR SMART: SM00181; EGF\_5.  
 DR PROSITE: PS00022; EGF\_1; UNKNOWN\_6.  
 DR PROSITE: PS01186; EGF\_2; 6.  
 KW EGF-like domain; glycoprotein.  
 SQ SEQUENCE 830 AA, 8743.1 MW, 81748.44 GCR209 G664.

Query Match 11.1%, Score 289, DB 4, Length 830;  
 Best Local Similarity 36.1%, Pred. No. 3.6e-19;  
 Matches 65, Conservative 14, Mismatches 79, Indels 22, Gaps 9;  
 QY 211 CEAKKWPENHLCTAMNNVWVHEDTSEICPPPMPTREKACELHIFGPTKEP--C 270  
 DB 210 CPGTGLNCSCT-CANCAACSPIDGSCCTPCWIGDTCCLPCDGFGLNCSHCDC 266  
 QY 271 QEGPKSYVFLPDY--GGSATGWKGLQNEACHPGFGYGDCKLPKSCNNGEMCDRF 328  
 DB 267 --KCKHNEPSPDGSSE 324  
 QY 329 -C-LCSPGMOGLOQR-----REGIPRMTPKIVD-----LPDHEVNSGKFNPKICKAS 373  
 DB 325 HQGECFLPWGSPPEE--GHTSTFGEDASTFLV--GHTSTFGEDASTFLV--GHTSTFGEDASTFLV 384  
 RESULT 9  
 Q43701  
 ID Q43701 PRELIMINARY: PRT: 830 AA.  
 AC Q43701;  
 DT 01-JUN-1998 (TrEMBLrel. 06, Created)  
 DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)  
 DE ACETYL LDL RECEPTOR PROTEIN.  
 GN ACETYL LDL RECEPTOR PROTEIN.  
 OS Homo sapiens (Human).  
 OC Eukaryota, Metazoa, Chordata, Craniata, Vertebrata, Euteleostomi,  
 OC Mammalia, Eutheria, Primates, Catarrhini, Hominoidea, Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC MEDLINE=98058897, PubMed-9495444;  
 RA Adachi H., Tsujimoto M., Arai H., Inoue K.;  
 RT "Expression cloning of a novel scavenger receptor from human  
 RT endothelial cells";  
 PL J. Biol. Chem. 272:31217-31220(1997).  
 DR EMBL: D86864; BAA24070.1; .  
 DR HSSP: P01180, 2BN2.  
 DR InterPro: IPR000561; EGF-like.  
 DR SMART: SM00181; EGF\_5.  
 DR PROSITE: PS00022; EGF\_1; UNKNOWN\_6.  
 DR PROSITE: PS01186; EGF\_2; 6.  
 KW EGF-like domain; glycoprotein; Receptor; Signal.  
 FT SIGNAL 1 15 POTENTIAL.  
 FT CHAIN 16 830 ACETYL LDL RECEPTOR.  
 SQ SEQUENCE 830 AA, 8743.3 MW, F56D9ELAA64L77 GRC64.

Query Match 11.1%, Score 289, DB 4, Length 830;  
 Best Local Similarity 36.1%, Pred. No. 3.6e-19;  
 Matches 65, Conservative 14, Mismatches 79, Indels 22, Gaps 9;  
 QY 211 CEAKKWPENHLCTAMNNVWVHEDTSEICPPPMPTREKACELHIFGPTKEP--C 270  
 DB 210 CPGTGLNCSCT-CANCAACSPIDGSCCTPCWIGDTCCLPCDGFGLNCSHCDC 266  
 QY 271 QEGPKSYVFLPDY--GGSATGWKGLQNEACHPGFGYGDCKLPKSCNNGEMCDRF 328  
 DB 267 --KCKHNEPSPDGSSE 324  
 QY 329 -C-LCSPGMOGLOQR-----REGIPRMTPKIVD-----LPDHEVNSGKFNPKICKAS 373  
 DB 325 HQGECFLPWGSPPEE--GHTSTFGEDASTFLV--GHTSTFGEDASTFLV--GHTSTFGEDASTFLV 384  
 RESULT 10  
 Q9Y151  
 ID Q9Y151 PRELIMINARY: PRT: 594 AA.  
 AC Q9Y151;  
 DT 01-NOV-1999 (TrEMBLrel. 12, Created)  
 DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)  
 DE KIAA0149 PROTEIN.  
 GN KIAA0149.  
 OS Homo sapiens (Human).  
 OC Eukaryota, Metazoa, Chordata, Craniata, Vertebrata, Euteleostomi,  
 OC Mammalia, Eutheria, Primates, Catarrhini, Hominoidea, Homo.  
 OX NCBI\_TaxID=9606,  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC MEDLINE=96127530, PubMed-8590280;  
 RA Nagase T., Seki N., Tanaka A., Ishikawa K., Nomura N.;  
 RT "Prediction of the coding sequences of unidentified human genes. IV.  
 RT the coding sequences of 40 new genes (K1AA0121-K1AA0160) deduced by  
 RT analysis of cDNA clones from human cell line KG-1";  
 RL DNA Res. 2:167-174(1995)  
 DR EMBL: D63483; BAA09770.1; .  
 DR HSSP: F01140, 2BN2.  
 DR InterPro: IPR000561; EGF-like.  
 DR SMART: SM00181; EGF\_5.  
 DR PROSITE: PS00022; EGF\_1; UNKNOWN\_6.  
 DR PROSITE: PS01186; EGF\_2; 6.  
 KW EGF-like domain; glycoprotein.  
 SQ SEQUENCE 830 AA, 8743.1 MW, 81748.44 GCR209 G664.





```

RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BRAIN;
RX MEDLINE=96360089; PubMed=9636030;
RA Nakayama M., Nishijima T., Nishise T., Nemura N., Seki N., Ohara O.;
RT "Identification of high-molecular-weight proteins with multiple RGF-
RT like motifs by motif-trap screening.";
RL Genomics 51:27-34(1998).
DR EMBL: AB011539; DAA42467.1; -
DR BSSP: P02468; IKD1;
DR InterPro: IPR000561; EGF-like;
DR InterPro: IPR002049; Laminin_RGF;
DR Pfam: PF00008; RGF; 2;
DR SMART: SM00180; RGF_Lam; 1;
DR SMART: SM00001; EGF_Like; 2;
DR PROSITE: PS00022; EGF_1; UNKNOWN_3;
DR PROSITE: PS01186; EGF_2; 2;
KW RGF-like domain; Glycoprotein;
FT NON_TER 1
SQ SEQUENCE 153 AA; 15579 MW; FF6E34E53803F45D CRC64;

```

```

Query Match          9.78; Score 253.5; DB 4; Length 153;
Best Local Similarity 34.88; Pred. No. le-16;
Matches 57; Conservative 12; Mismatches 72; Indels 23; Gaps 6;

QY 22 LVTACNNVCVHETDECTFPEEMETQEKAKELHTEHGTCKEPISQDECKSYVEGLF 282
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 2 LCV-CUHGAPCDPISGRLCLCFAGTGHGFCERGCEPSSGEGCIGQCD----CDGAGPCDP 56

QY 28 DYYGSCATGNGGCGNACHTGYGPDCKKRCSCNNGCMEDRFG-CGCSPOWGGLQCK 341
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 57 VTGLCLCPKRGSAICNLKCHFGQFQPSLTLTGAGGAGDPSVSSQGRVGVYMGPTC- 115

QY 342 REGIPRMTHKIVDLPDHIIEVNSGRNP-ICKASGMPLPTNEEMT 384
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 116 RRGGRRIPE-----NPSLAGDSAGTLPASSRPT 144

```

```

RESULT 15
QYBOXD
ID Q9H0X0 PRELIMINARY; PRI; 117 AA.
AC G9H0X0;
DE 01-JUN-2001 (Tremblrel. 17, Created)
DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)
DI 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE PROTEIN-TYROSINE KINASE TIE2 (FRAGMENT).
OS Sus scrofa (Pig).
OC Eukaryota, Metazoa, Chordata, Craniata, Vertebrata, Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
GX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21153163; PubMed=11230987;
RA Kim I., Moon S.O., Han C.Y., Pak Y.K., Moon S.K., Kim J.J., Koh G.Y.;
RT "The angiopoietin-tie2 system in coronary artery endothelium prevents
RT oxidized low-density lipoprotein-induced apoptosis.";
RL Cardiovasc. Res. 49:872-881(2001).
DR EMBL: AF251494; AAK18753.1; -
DR InterPro: IPR003961; FN_111;
DR Pfam: PF00041; FN3; 1;
DR SMART: SM00060; FN3; 1;
KW Kinase;
FT NON_TER 1
FT NON_TER 117
SQ SEQUENCE 117 AA; 13161 MW; 6C86HE99BF408B6 CRC64;

```

```

Query Match          8.58; Score 221; DB 6; Length 117;
Best Local Similarity 100.0%; Pred. No. le-13;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 431 GMVEKPFNLSVKVLKPLNAPNVIDIGHNFAVINISSEPYFG 472

```

```

DB 1 GMVEKPFNLSVKVLKPLNAPNVIDIGHNFAVINISSEPYFG 42

Search completed: September 23, 2002, 22:16:47
Job time: 123 sec

```





```

151 GluAspAlaValIleTyrLysAsnGlySerPheIleHisSerValProAr 167
|||||
599 CAACATGACGATGATACAAAATGCTTCCTTCATCCATTCAGTCCGCCG 648
167 qHisGluValProAspIleLeuGluValHisLeuProHisAlaGluProG 184
|||||
649 CATGAAGTATCTGATATCTGAATGAATGAATGAATGAATGAATGAATG 698
184 InAspAlaGlyValTyrSerAlaArgTyrIleGlyCysAsnLeuPheThr 200
|||||
699 AGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 748
201 SerAlaPheThrArgLeuIleValArgCysGluAlaGluLysTrpGlu 217
|||||
749 TCGGCCCTTCACACAGCGTGAATGATGATGATGATGATGATGATGATG 798
217 yProLeuCysAsnHisLeuLysThrAlaCysMetAsnAsnGlyValCysH 234
|||||
799 ACCTGAATGAAATCATCTCTGTATGATGATGATGATGATGATGATGATG 848
234 tSChAspPheThrGlyCysIleCysProProGlyPheMetClyArgThr 250
|||||
849 ATGAGATATGAGATGATGATGATGATGATGATGATGATGATGATGATG 898
251 CysGluLysAlaCysGluLeuHisThrPheGlyArgThrCysLysGluAr 267
|||||
899 TGTGACAAAGCTGCTGAAATGCTGACAGCTTGGCAGCAACTGTGTAAG 948
267 qCysSerGlyGluGluGlyCysLysSerTyrValPheCysLeuProAspP 284
|||||
949 GTGATGATGAAATGATGATGATGATGATGATGATGATGATGATGATG 998
284 rGlyCysSerCysAlaThrGlyTrpLysGlyLeuGlnCysAsnGlu 300
|||||
999 CTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1048
301 AlaCysHisProGlyPheTyrGlyProAspCysLysLeuArgCysSerCy 317
|||||
1049 GCAATGCAAGCCCTGGTCTTACAGGATGATGATGATGATGATGATGATG 1098
317 sAsnAsnGlyGluMetCysAspArgPheGlnIleLysLeuLysSerProG 334
|||||
1099 CAACATGAGGATGATGATGATGATGATGATGATGATGATGATGATGATG 1148
334 lYTrpGlnGlyLeuGlnCysGluArgGluGlyIleProArgMetThrPro 350
|||||
1149 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1198
351 LysIleValAspLeuProAspHisIleGluValAsnSerGlyLysPheAs 367
|||||
1199 AACAACTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1248
367 nProIleCysLysAlaSerGlyTrpProLeuProThrAsnGluGluMet 384
|||||
1249 TCCCATTTGTAAGATCTCTGATGATGATGATGATGATGATGATGATG 1298
384 hLeuValLysProAspGlyThrValIleHisProLysAspPheAsnHis 400
|||||
1299 CTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1348
401 ThrAspHisPheSerValAlaIlePheThrIleHisArgIleLeuProPr 417
|||||
1349 ACCGATCATTTTCATGACCAATATCAACATCCACCGGATCCCTCCGCC 1398
417 oAspSerGlyValTrpValCysSerValAsnThrValAlaGlyMetValG 434
|||||
1399 TGACTCAGGATTTGATGATGATGATGATGATGATGATGATGATGATG 1448
434 lLysProPheAsnIleSerValLysValLeuProLysProLeuAsnAla 450
|||||
1449 AAAGGCTTCACATTTCTGTTAAGTTCTTCCAAAGCCCTGAAAGCC 1498

```

451 ProAsnValIleAspThrGlyHisAsnPheAlaValIleAsnIleSerSe 467  
|||||

1499 CCAAACTGATTCACACATGACATAAATTTGCGTGGTTCATCAACATCAGGCTC 1548

467 rGluProIlePheGly 472  
|||||

1549 TGAATTTAATTGGG 1564

seq\_name: qb\_pr:HUMTEKRPK

seq\_documentation\_block:  
LOCUS HUMTEKRPK 4138 bp mRNA linear PRI 14 JAN 1995  
DEFINITION Homo sapiens receptor protein-tyrosine kinase (TEK) mRNA, complete cds.

ACCESSION 106139  
VERSION 106139.1 GI:292823  
KEYWORDS receptor protein-tyrosine kinase; transmembrane protein; tyrosine kinase.

SOURCE Homo sapiens  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 4138)  
AUTHORS Ziegler S F, Bird F A, Schneringer J A, Schooley K A, and Baum P R.

TITLE Molecular cloning and characterization of a novel receptor protein tyrosine kinase from human placenta  
JOURNAL Oncogene 8 (3), 663-670 (1993)

FEATURES  
Location/Qualifiers  
1..4138  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/tissue\_type="placenta"  
1..4138  
/gene="TEK"  
149..3523  
/gene="TEK"  
/codon\_start=1  
/product="receptor protein tyrosine kinase"  
/protein\_id="AAA61139.1"  
/db\_xref="GI:292823"  
/translation="MDSLASLVLGVSLLSGTVEGAMDLILINSLPLVSDAETSLTC  
TAGSRHEFTITGRDEALMNGHQPLEVITQVIREAKKVVKKREKASKINGAYEC  
EGRVGHAIRITMKMQOASPLPATITMTVDKDNVNTSFKKVLIKRDAVLYKNGS  
FIHSVPHVEVHDILVHLPHAQPDQACVYSARYIGGNLTSAPTRILIVRCAQKMG  
ECNHLCTACMNGVCHHDTGCTCPGPGMGKTCCKACELHITFCRTCKRCGQKGS  
YVFLCEFYGCSCATGKGLGCNEAGHGFYGEELKRCGNNNGEMTGLQGLDLSPG  
WQGLCEGEGIPBMTPKIVDLPDHIENVSKENPKICKASGWPLPTNEEMTLVKPGTV  
LHPKDFNHTDHSVAIFTIIRILPPDSGVWVSVNTVAGVKEPFIISVKVLKPLNA  
PNVITDGHNFVAINISSEPYFGDPIKSKLLYPVNVHYFAWHQIVTNEIVTNTLE  
PRTFYEICVQIVRKGEGEGHGVVFTTASTGLPPKGLNLIPASQFTLNITWQPI  
FPSSEDDFYVEVFRSVOKSDOONIKVPGNLTISVLLNLIHPROYVYVHARVNIKAQGE  
WSEDLTAWTLDLPPPENIKISNTHSSAVISWTILDGYSSTSTIRYKVGQNEF  
QHDVTKNATILQYOLKGLEPETAYQVDFEANNIGSSNPAFSEHSLVTLPESSQPAD  
LGCGKMLLIALLASAGMTCTLLAFLLIOLKRVNVRMAQAFONVREEPVQFNS  
GTLALNKKRNKNPIITYPVLDNDIKFOIVTIGRNPQGVILKRIKDGIRMDAAIKR  
MKFYAKDDHRDFAGELEVLCKLGHPIITNLIGACEHGYLYLALIEYHGNLIDPL  
RKSRVLTDPAFANSTASTLSQGLIHPAIVARGMIYLSQKQIHRDIAARNILV  
GENYVAKIDAGLSRGQEVYKRTMGRLPVRMAIESLNSVLTINSQVMSYGVLLWE  
IVSLGTPYCGMTCAELYEKLPQYKLEKPLNIDDEVYDLMRQWREKYEYEPSTAQI  
LVSINMLERKTYVNTTLYEKTYAGILGSAEEAA"  
149..202  
/gene="TEK"  
203..3520  
/gene="a a domains: 746..747, transmembrane: 211..340,  
EGF-like repeats: 440..733, fibronectin type 3 repeats"  
/product="receptor protein-tyrosine kinase"  
BASE COUNT 1170 a 910 c 988 g 1070 t  
ORIGIN

## alignment\_scores:

Quality: 265.00 Length: 472  
 Ratio: 5.519 Gaps: 0  
 Percent Similarity: 100.000 Percent Identity: 100.000

## alignment\_block:

US-09-733-764-2\_copy\_1\_472 x HUMTERPPTK

Align seq 1/1 to: HUMTERPPTK from: 1 to: 4138

1 MetAspSerLeuAlaSerLeuValLeuCysGlyValSerLeuLeuLeuSe 17  
 149 ATGAGATCTTTAGCAGCTTAGTCTTGTGAGATAGCTGACCTTC 198  
 17 rGlyThrValGluGlyAlaMetAspLeuIleLeuIleAsnSerLeuProL 34  
 199 TGGAACTGTGCAAGGTCCTCAAGGACTTGAATTCATCAATTCCTACCTC 248  
 34 euValSerAspAlaGluThrSerLeuThrCysIleAlaSerGlyTrpArg 50  
 249 TGTATCTGATGCTGCAACATCTGCACTGCACTGCACTGCACTGCACTG 298  
 51 ProHisGluProIleThrIleGlyArgAspPheGluAlaLeuMetAsnGl 67  
 299 CCGATAGAGCCATACCTATAGAGAGGAGCTTTGAGGCTTTATGAGACCA 348  
 67 rHisGlnAspProLeuGluValThrGlnAspValThrArgGluTrpAlaL 84  
 349 GCACCAAGATCCGCTGCAAGTTACTCAAGATGTCACCCAGCAAGATGGGCTA 398  
 84 ySLysValValTrpLysArgGluLysAlaSerLysIleAsnGlyAlaTyr 100  
 399 AAAAGTTGTTTGGAGAGAGAGAAAGGCTAGTAAAGATCAATGGCTTAT 448  
 101 PheCysGluGlyArgValArgGlyGluAlaIleArgIleArgThrMetLy 117  
 449 TCTCTGTAAGAGGAGATTTAGAGAGAGAGAGAGAGAGAGAGAGAGAG 498  
 117 sMetArgGlnGlnAlaSerPheLeuProAlaThrLeuThrMetThrValA 134  
 499 GATGGCTCAACAAAGCTTCCTTCTTACCAGAGTACTTTAACTATCAGCTG 548  
 134 sPLysGlyAspAsnValAsnIleSerPheLysLysValLeuIleLysGlu 150  
 549 ACAAGGAGATCAAGCTGAACATATCTTTCAAAAGGATTCATTAAGAA 598  
 151 GluAspAlaValIleTyrLysAsnGlySerPheIleHisSerValProAr 167  
 599 GAAGATGAGTGAATTTACAAAATGGTTCCTTCATCCATTCAGTGGCCG 648  
 167 rHisGluValProAspIleLeuGlyValHisLeuLeuLeuLeuLeuLeu 184  
 649 CATGAAGTACTGATATTCAGAGATACACCTGCCCTCATGCTCAGCCCC 698  
 184 InAspAlaGlyValTyrSerAlaArgTyrIleGlyAsnLeuPheThr 200  
 699 AGGATGCTGGAGTCTACTCGCCAGGATATAGGAGGAAACCTTCTCAC 748  
 201 SerAlaPheThrArgLeuIleValArgCysGluAlaGlnLysTrpC 217  
 749 TGGAGCTTATAGAGATATAGAGATATAGAGATATAGAGATATAGAGAT 798  
 217 yProGluCysAsnHisLeuCysThrAlaCysMetAsnAsnGlyValCysH 234  
 799 ACCTGAATGCACCACTCTCTACTGCTCTATGATGACCAATGCTCTCGCC 848  
 234 IsGluAspThrGlyLysCysLeuPheProGlyPheMetGlyArgThr 250  
 849 ATGAAGATCTGGCAATGCAATTTGCCCTCTCGGCTTATGGAAGGAGC 898  
 251 CysGluLysAlaCysGluLeuPheIleThrPheGlyArgThrCysLysGluAr 267  
 899 TGTGAGAGAGTCTTCTAAATTAAGATTTTAAATTAAGATTTTAAATTAAG 948

267 rGlySerGlyGlnIleGlyCysLysSerIleValPheLeuLeuLeuLeuLeu 84  
 949 GTGAGTGGACACACAGGCACTGAACCTCTATGAGTCTCTCTCTCTCTCT 998  
 284 rOTyrGlyCysSerCysAlaThrGlyTrpLysGlyLeuGlyCysAsnGlu 300  
 999 CTATGAGTCT 1048  
 301 AlaCysHisProGlyPheTyrGlyProAspCysLysIleuArgCysSerCy 317  
 1049 GATGAGTCT 1098  
 317 sAsnAsnIleGlyMetCysAspArgPheIleGlyCysLeuLysSerProG 334  
 1099 CAACAAATGGGACATGCTGCACTGCTGCAAGGATGCTGCTGCTGCTGCT 1148  
 334 LyTrpGlnGlyLeuGlnCysGluArgGlyIleProArgMetIlePro 350  
 1149 GATGAGTCT 1198  
 351 LysIleValAspLeuProAspHisIleGluValAsnSerGlyLysPheAs 367  
 1199 AAGATAGTCT 1248  
 367 nProIleCysLysAlaSerGlyTrpProLeuProThrAsnGluIleMet 384  
 1249 TCCCATTTGCAACGCTTCTGCTGGCGCTACCTTACTTACTTACTTACTT 1298  
 384 hrLeuValLysProAspGlyThrValLeuHisProLysAspPheAsnHis 400  
 1299 GCTGCTGAGAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1348  
 401 ThrAspHisPheSerValAlaIlePheThrIleHisArgIleLeuProP 417  
 1349 ACGATATTTTCTAGTAGTATATTATCTCTCTCTCTCTCTCTCTCTCT 1398  
 417 oAsnSerGlyValTrpValCysSerValAsnThrValAlaGlyMetValG 434  
 1399 TGACCTGAGAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1448  
 434 LuLysProPheAsnIleSerValLysValLeuProLysProLeuAsnAla 450  
 1449 AAAAGTCT 1498  
 451 ProAsnValIleAspThrGlyHisAsnPheAlaValIleAsnIleSer 467  
 1499 CCAAACTGATTAATCTGAGATATAATTTTCTCTCTCTCTCTCTCTCT 1548  
 467 rGluProTyrPheGly 472  
 1549 TGACCTTACTTTGGG 1564

seq\_name: qb\_cm:BTIE2A

seq\_documentation\_block:

LOCUS BTIE2A  
 DEFINITION B.taurus Tie 2 mRNA.  
 ACCESSION X71424  
 VERSION X71424.1 GI:296577  
 KEYWORDS receptor tyrosine kinase.  
 SOURCE cow.

ORGANISM

Bos taurus

Eukaryota; Metazoa; Chordata; Craniata, Vertebrata, Euteleostomi;  
 Mammalia, Eutheria, Cetartiodactyla, Ruminantia, Pecora, Bovidae;  
 Bovidae; Bovinae; Bos.

PEPENSE 1 (bases 1 to 4625)

AUTHORS Sato, T.N.

TITLE Direct Submission

JOURNAL Submitted (18-MAR-1993) T.N. Sato, Roche Institute of Molecular

Biolog, 340 Kingsland St., Nutley, NJ 07110, USA

REFERENCE 2 (bases 1 to 4625)

AUTHORS Sato, T.N., Gub.Y., Kozak, C.A. and Audis, K.L.

TITLE tie-1 and tie-2 define another class of putative receptor tyrosine kinase genes expressed in early embryonic vascular system  
JOURNAL Proc. Natl Acad Sci U S A 56 (26) 9355-9358 (1959)  
MEDLINE 94022374  
REMARK Erratum: [[published erratum appears in Proc Natl Acad Sci U S A 1993 Dec 15;90(24):12056]]

FEATURES  
Source location/Qualifiers

1..4625  
/organism="Bos taurus"  
/db\_xref="taxon:9913"  
/cell\_type="endothelial cells"  
/tissue\_type="brain"  
/pv\_stap="9431"  
324..377  
/gene="tie 2"  
324..3701  
/gene="tie 2"  
/codon\_start=1  
/product="receptor tyrosine kinase"  
/protein\_id="CAA50555.1"  
/db\_xref="GI:296578"  
/db\_xref="SWISS-PROT:Q06807"  
/translation="MDGLAGLVGVSLLSATVGGAGGLILINSLPLVSAETSLT"  
TASGWRDPEPTIGRDPALMNQHDPLEVTQATRWAKKVVVKREKASKNGATFC  
KORVKGQAIRTRIPKMKDQASFIIPATITMTVDEENVNISFKKVLKEEDAVIYNKCS  
FTHSVPREVPDILLEVQVPAQDQGVSAARYIGGNLFTSAFTRLITVRCAQKWP  
ECNRICTAMNNNGICHTDTEGICPPGFMGTCEKACEPHTFGPTCKRCSPEPSKS  
FVGLPDPVAGSCATGKGLQCNATJGYGYGDKLRGCTNGEKCDREDFGLCSPG  
RVLQCTKEVWPRMTKIEDIPIHIEVNSKFNPICKASGWRPPANEETLVKPDQTV  
LRPKDNNHICHLVAFTLNKILPPDSCVWVCNVNYSQWVEKPNISVKVLKPLNA  
PKVIDTCNPAVILSSRYEGGPIKSKKLLIKVNVNHYEARHLOVNEIVTLNTE  
PRTEYKCVQVRGGRGHPGVRRFTTASIGLPPRGISLIPKSOITLITWQPI  
FESSEDVYEVSRPVSVMNSDQNIKVPNLTSLVNLNHPPEQYIVAPVNTKAG  
EMSDLTAMTSDVPPQENIKIFNITSSAVISWTILDGYSISAIIRYKVGKNE  
DOHDVTKNATIQYQGLLEPQTVQVDVFAENNTGSSNPTSSHELTLSESOAPA  
DIGCKMILLALIGSAGMIGLVLAPLIMLOIKRANVORHMAQAFONVREPAVQFN  
SCTIALNKAANDNPEGLYPVLDNDLKFQVQVQCNFQGVKARIKKIDELMDAALK  
PKKIYASKDNRDPAAGLIVLCKLGHPNINILGACHPGYLYLALFAPHCNLLDF  
LRKSLIEDPAPFAIANSTSLSSQGLHFAALVARHGMETSKGFTIRGLAARNII  
VENVAFATPFTSPQFVYVVKTKMPLPVMMALESNYSVITNSDWSYGVLLW  
EIVSLGFTPGMTCAEYLEKFLQYPLEKPLNIDQVYVQLMPPWRKPYEHPSPFAQ  
ILVSLNMLERKRYVNTILYKFTIYAGIDCSAPEAA"

324..3701  
/gene="tie 2"  
378..3698  
/gene="tie 2"  
/product="receptor tyrosine kinase"  
BASE COUNT 1292 a 1064 c 1108 g 1162 t  
ORIGIN

alignment\_scores:  
Quality: 24.5/00 Length: 472  
Ratio: 5.259 Gaps: 0  
Percent Similarity: 98.093 Percent Identity: 93.220

alignment\_block:  
US-09-733-764\_2\_COPY\_1\_472 a BT1E2A ..

Align seq 1/1 to: BT1E2A from: 1 to: 4625

1 MetAspSerLeuAlaSerLeuValLeuGlyValSerLeuLeuSeu 17

|||||

424 ATGGATCTTTAGCGGCTTATGCTCTGCGAGTCAGCTTCTCTTC 373

|||||

17 rGlyThrValGluGlyAlaMetAspLeuLeuLeuLeuLeuLeuLeu 34

|||||

374 TGAACAGTGGACGGTCCATGCACTGATGATGATGATGATGATGAT 423

|||||

34 euValSerAspAlaGluThrSerLeuThrCysLeuAlaSerGlyTrpArg 50

|||||

424 TTGTAICTGATGCTGAGACATCCCTACATGATGATGATGATGATGAT 473

|||||

51 ProHisGluProIleThrIleGlyAlaGlyAspPheGluAlaIleuMetAsnGlu 67

|||||

474 GATCAT 523

|||||

67 nHisGluAspProLeuGluValThrGlnAspValThrArgGluTrpAlaAla 84

|||||

524 GCACTGACAT 573

|||||

84 ystLysValValIleTrpLysAspGlnLysAlaSerLysIleAsnGlyAlaTrp 100

|||||

574 AAAAGTTGTTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 623

|||||

101 PheCysLeuGlyArgValArgGlyGlyAlaIleCysPheIleCysArgMetIly 117

|||||

624 TTTCTGTAAG 673

|||||

117 sMetArgGlnGlnAlaSerPheLeuProAlaThrLeuThrMetThrValAla 134

|||||

674 GATGGGTCAACAAAGGTTCTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 723

|||||

134 sPlysGlyAspAsnValAsnIleSerPheLysLysValLeuLeuLysGlu 150

|||||

724 AAG 773

|||||

151 GluAspAlaValIleCysLysAsnGlySerPheIleHisSerValProArg 167

|||||

774 GAAAG 823

|||||

167 qHisGluValProAspIleLeuGluValHisLeuProHisAlaGlnProG 184

|||||

824 GATAG 873

|||||

184 LeuAspAlaGlyValIleSerAlaArgTyrIleGlyLysAsnLeuSerPheTrp 200

|||||

874 AGATATGAG 923

|||||

201 SerAlaPheThrArgGluIleValArgCysGluAlaGlnLysIlePro 217

|||||

924 TAAAG 973

|||||

217 yProGluCysAsnAlaLeuCysThrAlaCysMetAsnGlyValCysH 234

|||||

974 ACCTGATGTAACCGCATCTGCACCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1023

|||||

234 IscLysAspThrGlyGlyCysIleCysProGlyPheMetGlyArgThr 250

|||||

1024 ATGAG 1073

|||||

251 CysCysLysAlaCysGlnLysHisThrPheGlyArgThrCysLysCysLys 267

|||||

1074 TGTGAG 1123

|||||

267 yCysSerGlyValGluGlyLysSerTyrValPheCysLeuTrpAspP 284

|||||

1124 GTGATGAG 1173

|||||

284 rGlyGlyCysSerCysAlaThrGlyTrpLysGlyLeuGlnCysGlnGlu 300

|||||

1174 CCTATGGGTGTTCTCTGTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1223

|||||

301 AlaCysHisProGlyPheThrCysGlyProAspCysLysLeuArgCysSerCys 317

|||||

1224 GATAG 1273

|||||

317 sAsnAsnGlyGluMetCysAspArgPheGlnGlyCysLeuCysSerProG 344

|||||

1274 GATCAATGGGAG 1323

|||||

334 LysTrpGlnGlyLeuGlnCysGluArgGluGlyIleProArgMetThrPro 350

|||||

1324 GATCAATGGGAG 1373

|||||

351 LysIleValAspLeuProAspHisIleGluValAsnSerGlyLysPheAs 367

|||||



[illegible]

```

451 ProAsnValIleAspThrGlyHisAsnPheIaValIleAsnIleSerSe 467
1437 CCAAACTGATTCACACACGACGACAAIAACTTTCCTATCAAAIACGCTC 1486
467 rGluProTyPheGly 472
1487 TGACCTTACTTTGGG 1502
seq_name: qb_pat:E08401
seq_documentation_block:
  Locus E08401
  DEFINITION cDNA of mouse TIE-2 receptor..
  ACCESSION F08401
  VERSION F08401.1 GI:2176518
  KEYWORDS JP 194315382-A/1.
  SOURCE Mus sp..
  ORGANISM Mus sp..
  Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae;
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata;
  1 (bases 1 to 3760)
  Murayama,Y., Suda,T., Iwama,A., Yasuhara,K. and Masu
  TIE-2 RECEPTOR AND DNA CODING THE SAME RECEPTOR
  Patent: JP 1994-15382-A 1 15-NOV-1994;
  SUDA TOSHIO, YAMANOUCHI PHARMACEUT CO LTD
  CS Mus sp. (mouse)
  PN JP 1994315382-A/1
  PD 15-NOV-1994
  PF 06-MAY-1994 JP 1993129912
  FI M'PAYAMA YUKO, SUDA TOSHIO, IWAMA ATSUSHI, YASU
  MASYASU YASUHIKO
  CC strandedness: Double;
  CC topology: linear;
  CC feature is identified by similarity;
  FH Key Location/Qualifiers
  FT Source 1..3760
  FT /organism='Mus sp.'
  FT /cell_line='DA-1'
  FT sig_peptide 152..211
  FT /mat_peptide 212..4523
  FT /product='mouse TIE-2 receptor'
  FT Location/Qualifiers
  1..3760
  /db_xref='taxon:10095'
  BASE COUNT 1024 a 876 c 970 g 890 t
  ORIGIN
alignment_scores:
  Quality: 2407.00 Length: 472
  Ratio: 5.233 Gaps: 0
  Percent Similarity: 97.458 Percent Identity: 91.949
alignment_block:
US-05-7335-764-2_copy_1_472_x_E08401
Align seq 1/1 to: E08401 from: 1 to: 3760
1 MetAspSerIleAlaSerLeuValIleuCysGlyValSerIleuLeuSe 17
|||||
152 APlAAAPPIIAAGNGGPTTACTTCTCTCTACAGACAGTCTCTCTTA 201
17 rGlyThrValGluGlyAlaMetAspLeuIleuLeuAsnSerLeuProl 44
|||||
202 GAGAGTATGACAGAGAGAGGATACAGAGAGAGAGAGAGAGAGAGAG 251
34 euValSerAspAlaGluThrSerLeuGluThrGlyTicAlaSerGlyTyrArg 50
|||||
252 TTATGTTGATGAGGAGAGAAATATATTTAAATATATATATATATATGATGACAC 401

```







```

334  lyTrpGlnGlyLeuGlnCysGlnAurGlnGlyIleProArqMetThrPro 350
1124 CATGGCAAGGGGTCACAGTGTGTGACAAAGAGAGGACAGGCTAAGCGATCAGCTCA 1173
351  LysIleValAspLeuProAspHisIleGluValAsnSerGlyLysPheAs 367
1174 CAGATAGAGCAATTTGTAAGATTAATATGAGGTAAATATGTGAAAAATTTAA 1223
367  nproIleCysLysAlaSerGlyTrpProIleuproThrAsnGluGluMet 384
1224 CCGCAATCTCAAAATCTGCTGGTGGCAATATATAGTGAAGGAATATGA 1273
384  hrLeuValLysProAspGlyThrValLeuHisProLysAspPheAsnHis 400
1274 CCGTAGTGAAGGCTAGATGGAGAGTGTCTCCAAACAAAGAGACTTCAACAT 1323
401  ThrAspHisPheSerValAlaIlePheThrIleHisArgIleLeuProPr 417
1324 ACAGATAGTGTATATATAGTGGTATATATCTGTCAAAGAGAGTGTATATCTG 1473
417  cAspSerGlyValTrpValCysSerValAsnThrValAlaGlyMetValC 434
1374 TGACTTAGAGATTTGGATTTGAGTGTGAACATATATGATGGATGGTGG 1423
434  luLysProPheAsnIleSerValLysValLeuProLysProLeuAsnAla 450
1424 AAAAGCCCTTCAACATTTCGTCAAAGTCTTCCACAGGCCCTTCGACGCC 1473
451  ProAsnValIleAspThrGlyHisAsnPheAlaValIleAsnIleSerSe 467
1474 CCAAAATGTGATTCAGACAGAGCAATAAATTGTGTATCATCAATAATCACTC 1523
467  rGluProTyrPheGly 472
1524 TTAGGCTTATTTTGG 1539
seq_name: qb_pat: I71113
seq_documentation_block:
LOCUS       I71113                               4175 bp          linear
DEFINITION  Sequence 1 from patent US 5681714.
ACCESSION   I71113
VERSION     I71113.1  GI:3007248
KEYWORDS    .
SOURCE      Unknown.
            ORGANISM: Unknown.
            Unclassified.
REFERENCE   1 (bases 1 to 4175)
AUTHORS    Breilman,M.L. deceased, Rossant,J., Dumont,D.J. and
TITLE      Nucleic acid encoding tek receptor tyrosine kinase
JOURNAL    Patent: US 5681714-A 1:28-OCT-1997;
FEATURES   location/Qualifiers
            source
            1..4175
            /organism="unknown"
BASE COUNT  1175 a    942 c    1047 g    1011 t
ORIGIN

```

```

alignment_scores:
  quality: 2407.09      Length: 472
  Ratio: 5.233          Gaps: 0
  Percent Similarity: 97.458  Percent identity: 91.949

alignment_block:
  US-09-733-764-2_COPY_1_472 x 171113  ..

Align seg 1/1  to: 171113  from: 1  to: 4175

1 MetAspSerLeuAlaSerLeuValLeuCysGlyValSerLeuLeuLeuSe 17
|||||
124 ATGGATTTTAAAGAGGTTAAATTTTATGGATATAGATATATTCCTTA 173

```



1024	GPATATGTAATLGGTAACTAGAGAAATATGTAAAGTAAAGTAACTAG	1073
317	SASNAASnGlyGluMetCysAspArgPheGlnGlyCysLeuCysSerProG	334
1074	TAACCAACAACAGAAATAGTGAAGGATCCAAAGCAAGATCGGCTCAAG	1123
334	LyTrpGlnGlyLeuGlnCysGluArgGluIleProArgMetThrPro	350
1124	CATGGCAAGGCTTGTATGTCAGAAAGAGGTAAGGTAAAGCATGACATCCA	1173
351	lysIleValAspLeuProAspHisLeuIleValAsnSerGlyIysPheAs	367
1174	CAGATAGAGGATTTGGTAGATCATATTGAGTAAACAGTGGAAATTTTAA	1223
367	nProIleCysIysAlaSerGlyTrpProLeuProThrAsnGluGluMetI	384
1224	CCCCATCTCCAAAGCCCTCGCGTCCGCACACCTACCTACTGCACAAAATGA	1273
384	hrLeuValLysProAspGlyThrValLeuIleHisProLysAspPheAsnIle	400
1274	CCCTAGTCAAGCAAGATGGTACATGGTCCAAAGCAAAAGATCAACATAT	1323
401	ThrAspHisPheSerValAlaIlePheThrIleHisAlaGlyLeuProPr	417
1324	ACAAATATTTTCTATGATATATATATATATATATATATATATATATAT	1373
417	oAspSerGlyValTrpValCysSerValAsnThrValAlaGlyMetValG	434
1374	TGCACTCAGTAGTCTGGGCTGGCAATGTCAAATATATATATATATATAT	1423
434	IuLysProPheAsnIleSerValLysValLeuProLysProLeuAsnAla	450
1424	AAAAAGCCCTTCAACATTTCCGTCAAGTCTTCACAGAGCCCCATCAGGCC	1473
451	ProAsnValIleAspIhrGlyHisAsnPhcAlaValIleAsnIleSerSe	467
1474	CCAAATATGATTCATATATGATATATATATATATATATATATATATAT	1523
467	gGluProIlyrPhcGly	472
524	TCAGCCTTACTTTGGG	1539

seq_documentation_block:			
LOCUS	171115	4176 bp	linear
		DNA	linear
			PAR 03-APR-1998

seq_documentation_block:			
LOCUS	171115	4176 bp	linear
		DNA	linear
			PAR 03-APR-1998

VERSION 1.0  
KEYWORDS

ORGANISM Unknown.

REFERENCE 1 (bases 1 to 4176)

TITLE Nucleic acid encoding tek receptor tyrosine kinase

FEATURES  
Location/Qualifiers

	1978	1979	1980	1981	1982	1983	1984	1985	1986	1987	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024	2025	2026	2027	2028	2029	2030	2031	2032	2033	2034	2035	2036	2037	2038	2039	2040	2041	2042	2043	2044	2045	2046	2047	2048	2049	2050	2051	2052	2053	2054	2055	2056	2057	2058	2059	2060	2061	2062	2063	2064	2065	2066	2067	2068	2069	2070	2071	2072	2073	2074	2075	2076	2077	2078	2079	2080	2081	2082	2083	2084	2085	2086	2087	2088	2089	2090	2091	2092	2093	2094	2095	2096	2097	2098	2099	2100	2101	2102	2103	2104	2105	2106	2107	2108	2109	2110	2111	2112	2113	2114	2115	2116	2117	2118	2119	2120	2121	2122	2123	2124	2125	2126	2127	2128	2129	2130	2131	2132	2133	2134	2135	2136	2137	2138	2139	2140	2141	2142	2143	2144	2145	2146	2147	2148	2149	2150	2151	2152	2153	2154	2155	2156	2157	2158	2159	2160	2161	2162	2163	2164	2165	2166	2167	2168	2169	2170	2171	2172	2173	2174	2175	2176	2177	2178	2179	2180	2181	2182	2183	2184	2185	2186	2187	2188	2189	2190	2191	2192	2193	2194	2195	2196	2197	2198	2199	2200	2201	2202	2203	2204	2205	2206	2207	2208	2209	2210	2211	2212	2213	2214	2215	2216	2217	2218	2219	2220	2221	2222	2223	2224	2225	2226	2227	2228	2229	2230	2231	2232	2233	2234	2235	2236	2237	2238	2239	2240	2241	2242	2243	2244	2245	2246	2247	2248	2249	2250	2251	2252	2253	2254	2255	2256	2257	2258	2259	2260	2261	2262	2263	2264	2265	2266	2267	2268	2269	2270	2271	2272	2273	2274	2275	2276	2277	2278	2279	2280	2281	2282	2283	2284	2285	2286	2287	2288	2289	2290	2291	2292	2293	2294	2295	2296	2297	2298	2299	2300	2301	2302	2303	2304	2305	2306	2307	2308	2309	2310	2311	2312	2313	2314	2315	2316	2317	2318	2319	2320	2321	2322	2323	2324	2325	2326	2327	2328	2329	2330	2331	2332	2333	2334	2335	2336	2337	2338	2339	2340	2341	2342	2343	2344	2345	2346	2347	2348	2349	2350	2351	2352	2353	2354	2355	2356	2357	2358	2359	2360	2361	2362	2363	2364	2365	2366	2367	2368	2369	2370	2371	2372	2373	2374	2375	2376	2377	2378	2379	2380	2381	2382	2383	2384	2385	2386	2387	2388	2389	2390	2391	2392	2393	2394	2395	2396	2397	2398	2399	2400	2401	2402	2403	2404	2405	2406	2407	2408	2409	2410	2411	2412	2413	2414	2415	2416	2417	2418	2419	2420	2421	2422	2423	2424	2425	2426	2427	2428	2429	2430	2431</
--	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	--------

BASE COUNT	1176 a	939 c	1048 q	1013 t
ORIGIN				

Leaflet: 407.00  
Leaflet: 472

Percent Similarity: 97.458 Percent Identity: 91.949

alignment\_block:

Align seq 1/1 to: 171115 from: 1 to: 4176

```

1 MetAspSerLeuAlaSerLeuValLeuCysGlyValSerLeuLeuLeuSe 17
|||||
124 ATGGACTCTTTAGTGGGTTATTCTCTGTGGAGTCAGCTGCTCTTTA 173
|||||
17 rGlyThrValGluGlyAlaMetAspLeuLeuLeuLeuLeuLeuLeu 34
|||||
174 TCGACTTATGAAAGGCGGCAAGCAAGTCACTTCAATTCCTCACTC 223
|||||
34 euValSerAspAlaGluThrSerLeuThrCysIleAlaSerGlyTrpArg 50
|||||
224 TTGTGCTGATGCGGAACATCCCTACCTGCAATTCCTGCGGGCCAC 273
|||||
51 proHisGluProIleThrIleGlyArgAspPheGluAlaLeuMetAsnG 67
|||||
274 GCGCATGAGGCGATACCATAGAGAGAGATTTGAAGGCTTAATGAAACA 323
|||||
67 nHisGluAspProLeuGluValThrGlnAspValThrArgGluTrpAlaI 84
|||||
424 GCAACCAAGATCCACTGCGAGGTTACTCAAGATGTCACCAAGAAATGGCGA 373
|||||
84 yLysValValTrpLysArgGluLysAlaSerLysIleAsnGlyAlaTy 100
|||||
374 AAAAGTTGTTGGCAAGACAGAAAGGCGAGTAAAGATTAAAGGCTTAT 423
|||||
101 PheCysGluGlyArgValArgGlyGluAlaIleArgIleArgThrMetLy 117
|||||
424 TTCTGTGAAGTGGAGTTCTGAGAGAGAGATATTAAGATACAGAGATGAA 473
|||||
117 sMetArgGluGlnAlaSerPheLeuProAlaThrLeuThrMetThrValA 134
|||||
474 GATGGCTTACAAAGATCTCTCTACCTGCTACTTTAACTAAGCGGTGG 523
|||||
134 sPlyGlyAspAsnValAsnIleSerPheLysLysValLeuIleLysGlu 150
|||||
524 ACAGGGGACATAATGUGAACAATCTTCACAAAAGGCTGTAAATTAAGAA 573
|||||
151 GluAspAlaValIleTyrLysAsnGlySerPheIleHisSerValProAr 167
|||||
574 GAAGATGCTACTATTACAAAATAGCTCTTCATCCACTCAGTGGCGCG 623
|||||
167 gHisGluValProAspIleLeuGluValHisLeuProHisAlaGlnPro 184
|||||
624 GCAAGAGATACGATATTTAATAAGTTCATCTGCGCATGCTCAGCGCC 673
|||||
184 InAspAlaGlyValTyrSerAlaArgTyrIleGlyLysLeuPheThr 200
|||||
674 AGCATGCTGCTGTACTGCGGCGCAGCTACATAGGAGGAACCTGTTCC 723
|||||
201 SerAlaPheThrArgLeuIleValArgArgCysGluAlaGlnLysTrpG 217
|||||
724 TCAGCTTACAGAGTGTATTATTGGNATATGTAAATTCATAGTGGGG 773
|||||
217 yProGluCysAsnHisLeuCysThrAlaCysMetAsnGlyValLysH 234
|||||
774 GCGCGCATGTAAGCGGCTGTACTTCTACTTGGCAAGCAATGGAGTCC 823
|||||
234 IsdAspThrGlyGlyCysIleCysProProGlyPheMetGlnArgThr 250
|||||
824 ATCAACATATGCGGAATGATATGGGCTGCTGCTGATGGGCAACA 873
|||||
251 CysLeuLysAlaCysLeuLeuHisSerPheCysArgThrCysLysLeu 267
|||||
874 TTGTGAAGAGCTGTGAGTGGCAGACATTTCGAGAGCTTAAAGAG 923
|||||
267 qYsserGlyGluGluGlyLysSerSerValPheCysLeuProAspP 284
|||||
924 GTATATGCAATACAGAGTAAGAGTATGATGTTGATGATGATGATG 973
|||||
284 rGlyThrValGluGlyAlaMetAspLeuLeuLeuLeuLeuLeuLeu 300
|||||
974 CTACAGGCTTCTGCGGATACAGCTGACAGAGGAGGAGGAGGAGAA 1023

```

```

301 AlaCysHisProGlyPheThrGlyProAspCysLysLeuGlyCysSerCy 417
|||||
1024 GCAAGGATACCTGGTTCAGTACAGACAGAGTGAAGCTCAGTACAGCTG 1073
|||||
317 sAsnAsnGlyGluMetCysAspArgPheGlnGlyCysLeuCysSerPro 434
|||||
1074 TATCAATCAAGACATATCTCATCTCCACGAGTGGCTCTCTCTCTCT 1123
|||||
334 LyTrpGlnGlyLeuGluCysGluArgGluGlyIleProArgMetThrPro 450
|||||
1124 GATGATGAGAGTGTAGTGTGAGAAAGAGAGAGAGAGAGAGAGAGAG 1173
|||||
351 IysIleValAspLeuProAspHisIleGluValAsnSerGlyLysPheAs 467
|||||
1174 CAAATAGAGATTTCAGATTAATTAATTGAAATTAATTAATTAATTAAT 1223
|||||
367 nProLysLysAlaSerGlyTrpProLeuProThrAsnGluIleMet 484
|||||
1224 CCGCATCTGTAAGGCTCTGGGTGGCTCACTACTACTACTACTACTACT 1273
|||||
384 hLeuValLysProAspGlyThrValLeuHisProLysAspPheAsnHis 400
|||||
1274 CCAATATGAGAGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 443
|||||
401 ThrAspHisPheSerValAlaIlePheThrIleHisArgIleLeuPro 417
|||||
1324 ACAGATGCTTCTGAGTGGCAATATGATGATGATGATGATGATGATG 1473
|||||
417 oAspSerGlyValTrpValCysSerValAspThrValAlaGlyMetVal 444
|||||
1374 TGACTCAGAGCTGAGCTGTAGTGTGAGAGAGAGAGAGAGAGAGAG 1423
|||||
434 LuLysProPheAsnIleSerValLysValLeuProLysProLeuAsnAla 450
|||||
1424 AAAAGCTTTCAGATTCGTCGCAAGTCTTCAGAGAGAGAGAGAGAG 1473
|||||
451 ProAsnValIleAspIleHisAsnPheAlaValIleAsnIleSerSe 467
|||||
1474 CCAATGATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1523
|||||
467 rGluProTyrPheGly 472
|||||
1524 TCGAGCTTACTTTGGG 1539

```

seq\_name: gb\_ro:MMTEK

seq\_documentation\_block:

LOCUS MMTEK 4176 bp mRNA linear REF 04-DEC-1994

DEFINITION M. musculus mRNA for tek.

ACCESSION X67553.3 S40311

VERSION X67553.3 GI:297158

KEYWORDS tek gene; Tyrosine kinase.

SOURCE house mouse.

ORGANISM

Mus musculus

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

1 (bases 2587 to 4176)

Breitman, M.L.

tek, a novel tyrosine kinase gene located on mouse chromosome 4, is

expressed in endothelial cells and their presumptive precursors

Oncogene 7 (8), 1471-1480 (1992)

9234855

2 (bases 1 to 4176)

Dumont, P.J.

Direct Submission

Submitted (26-APR-1994) D.J. Dumont, Samuel Lunenfeld Res.

Institute, Mount Sinai Hospital, 600 University Ave, Toronto,

Ontario, M5G 1X5, CANADA

3 (bases 1 to 4176)

Dumont, D.J., Gradwohl, G.J., Foug, G.H., Auerbach, R. and

Breitman, M.L.

The endothelial-specific receptor tyrosine kinase, tek, is a member

of a new subfamily of receptors  
 JOURNAL  
 MEDLINE  
 93241731  
 COMMENT  
 The endothelial-specific receptor tyrosine kinase, Tek, is a member  
 of new subfamily of receptors  
 Related sequence: X67553.

## FEATURES

source  
 1. 4176  
 /organism="Mus musculus"  
 /strain="Cd-1"  
 /db\_xref="taxon:10090"  
 /chromosome="4"  
 /tissue\_type="embryonic heart."  
 /dev\_stage="day 12.5"  
 124. 3492  
 /gene="tek"  
 124. 3492  
 /gene="Tek"  
 /codon\_start=1  
 /product="receptor tyrosine kinase"  
 /protein\_id="CAA47857.1"  
 /db\_xref="GI:297159"  
 /db\_xref="MGI:98664"  
 /db\_xref="SWISS-PROT:Q02858"  
 /translation="MDSLGLVLCVSLLLYGVVSGAMLLILNSLPLVSAETSLIC  
 IASGHPHEPTIGRDFALMNDHQDLEVTQVTRWAKKVMKREKASKINGAYFC  
 EGRVQQAIRIRMKMOOASLPATLMTVDGDNVNISEKKVLIKEDAVIYKNGS  
 FTHSVRRHPDILVHILPHAQIDAGVYSAHYIGENLFTSAFLHIVRRFAKWKGP  
 DCSRPCTICKNVCHEHTGECICPPGFMGRCEKAFPHPTGRTCKRCQSPGCKS  
 VYFCLPDYGVSCATGRLQCEACPSGVGPNKLPCHCTNEEDICPFQGCILSQG  
 WGLQCEKBRKMPQLLEDLPHLLEVANSKFNPCIKASGMLPTESEMTLVKPDGTV  
 LQPDENYTDRESVAIFVNRVLPEDNSVWVTSYNTVA-SMVEKPNISVKKVLPPLHA  
 PNVIDGHNFALINISSEHYFGDGPISKKKLPYKPVNOAWKYIETNREIPTNLYEPR  
 TDYELQVQARPCQEGHGPVRRPFTTACIGLPPPGISLLPKSQALNLTWQPIFT  
 NSHDFYVHVEERSQVTSQDQNIKVPGNLTSVLISNIVPREQYTVRARNVTKAOGEW  
 SEELRAWTLSDILPQENIKISNITDSTAMVSWTIVDYSTISSTIIIPYKVGKNEDW  
 HDVKIKNAITVOYOLKGLPETTVYVDIEAENNIGSSNPAESHELRLTLPSPGSADL  
 GAEKMLLITAIASASMTQITVLLAFILMLLKPANVLPMAAPLNREEPAPVFNSTG  
 LA-NRKAANNDDITIVYVLDNDIRFQDVIGECNFCVILKAKIKDDQIRMADAIKRMK  
 EVASDDHDFPAGRIKVLCKTGHPNINILGACHRGVYVIAIYAPHGLNLPK  
 SVLEIDPAIALANSQQLLHFAADVARGMDYLSQKQFIHRDLAARNILVGE  
 NYIAKIDRLSPQEVYVKTKMPLPVRMAIESINTSVYTNDSVMSYGVLLWEIV  
 SLGTYGCMTCALVEKLPQGYRLEKPLNDDVDYDLMRQWRKPYRPSFAQILV  
 SLNRMLEERKTVNTLLYEKTYAGIDCSAEAA"

BASE COUNT 1176 a 339 c 1048 g 1013 t

ORIGIN

alignment\_scores:  
 Quality: 24.00 Length: 472  
 Ratio: 5.233 Gaps: 0  
 Percent Similarity: 97.458 Percent identity: 91.949

## alignment\_block:

US-09-733-764-2\_COPY\_1\_472 x MMTEK

Align seg 1/1 to: MMTEK from: 1 to: 4176

1 MetAspSerLeuAlaSerLeuValLeuCysGlyValSerLeuLeuLeuSe 17  
 124 ATGGAATTTTAGGAGATTTATTTCTGTGTGTGATGATCAATTCCTACCTC 173  
 17 rGlyThrValGluGlyAlaMetAspLeuLeuLeuLeuLeuLeuLeuLeu 34  
 174 TCGAGTAGTACAGGCGGCAATGAGACCTGATCTTCATCAATTCCTACCTC 223  
 34 euValSerAspAlaGluThrSerLeuThrCysIleAlaSerGlyTrpArg 50  
 224 TTGCTCTCTGATGCCGGAACATCCCTCACCCTGCAATGCCCTCTCGGTCAC 273  
 51 ProHisGluProIleThrIleGlyArgAspPheGluAlaIleuMetAsnG 67  
 274 GCGCATAGCGCATATATATATATATATATATATATATATATATATATAT 323

67 nHisGlnAspProLeuGluValThrGlnAspValThrArgGluTrpAlaL 84  
 324 GAAAGAAAT 373  
 84 ySLysValValTrpLysArgGluLysAlaSerLysIleAsnGlyAlaTrp 100  
 374 AAAAAGTTTCTTTTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 423  
 101 PhcCysGluGlyArgValArgGlyGluAlaIleArgIleArgThrMetIly 117  
 424 TCTGTGAAGGTCAGTTCAGGACAGGCTATAAGGATACAGGACAAAGAA 473  
 117 sMetArgGlnGlnAlaSerPheLeuProAlaThrLeuThrMetThrValA 134  
 474 GATCGCTCAACAAAGCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 523  
 134 spLysGlyAspAsnValAsnIleSerPheLysLysValLeuIleLysGlu 150  
 524 ACAGGAGAGAT 573  
 151 GluAspAlaValIleTyrLysAsnGlySerPheIleHisSerValProAr 167  
 574 GAAGATCTAT 623  
 167 qHisGluValProAspIleLeuGluValHisLeuProHisAlaGlnProG 184  
 624 GATGAAGAGATATATATATATATATATATATATATATATATATATAT 673  
 184 InAspAlaGlyValTyrSerAlaArgTyrIleGlyGlyAsnLeuPheThr 200  
 674 AGAT 723  
 201 SerAlaPheThrArgLeuIleValArgArgCysGluAlaGlyLysTrpG 217  
 724 TCAGCCTTCACCAAGGCTGATTTTTCAGAGATATGAGATCTCAAGCTGGG 773  
 217 yProGluCysAsnIleLeuCysThrAlaCysMetAsnAsnGlyValCysH 234  
 774 GCGGAT 823  
 234 iGluAspThrGlyGluCysIleCysProProGlyPheMetGlyArgTrp 250  
 824 ATGAAAT 873  
 251 CysGluLysAlaCysLysLeuHisThrPheGlyArgThrCysLysGluAr 267  
 874 TATTAATAAAAT 923  
 267 qCysSerGlyGluGluGlyCysLysSerTyrValPheCysLeuProAsp 284  
 924 GTGTAGTGGACATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 973  
 284 rGlyThrValGlySerCysAlaThrGlyTrpPheGlyGlyGlyGlyGlyG 309  
 974 CTATACAGGAT 1023  
 301 AlaCysHisProGlyPheTyrGlyProAspCysLysLysLeuArgCysSer 317  
 1024 GATGAGCAT 1073  
 317 sAsnAsnGlyGluMetCysAspArgPheGlnGlyCysLeuCysSerProG 334  
 1074 TATCAATAGAGACATATATATATATATATATATATATATATATATATAT 1123  
 334 lYTrpGlnGlyLeuGlnCysGluArgGlyIleProArgMetThrPro 350  
 1124 GATGAG 1173  
 351 lYsIleValAspLeuProAspHisIleGluValAsnSerGlyLysPheAs 367  
 1174 CAGATATAGGATATATATATATATATATATATATATATATATATATAT 1223











```

940 TCAGCGCTGTCGCGACGCGTAAATACGCGCTGGCTGTATACACAAAGCTGTG 989
116 eFCysAsnAsnGlyGlyMetCysAspArgPheGlyCysLeuCysSer 332
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
990 TGTGTGCAAAAGAGAG TGGATGATGATGATGATGATGATGATGATGATG 1046
433 ProGlyThrGlyCysGlyCysGlyCysGlyCysGlyCysGlyCysGly 349
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1047 ...GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1074
449 rProLysIleValAspLeuProAspHisIleGluValAsnSerGly...L 365
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1075 CCGCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1121
465 rPheAsnProIleCysGlyAlaSerGlyTyrProLeuProThrAsnGlu 381
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1122 AGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1171
482 GluMetThrLeuValIleAspGlyThrValIleGluHisProLysAsp 398
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1172 GATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1221
498 eAsnIleThrAspHisPheSerValAlaIlePheThrIleHisAcqIleL 415
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1222 ACACACACACACACACACACACACACACACACACACACACACACACACAC 1271
415 euProProAspSerGlyValIleValCysSerValAsnThrValAlaGly 431
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1272 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1321
432 MetValGluIleLysProGlyAsnIleSerValIleGlyValIleGly 448
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1422 CAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1371
448 uAsnAlaProAsnValIleAspThrGlyHisAsnPheAlaValIleAsnI 465
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1472 CAATGCTGCGACGACGACGACGACGACGACGACGACGACGACGACGACG 1421
465 leSerSerGluProTyrPheGly 472
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1422 TCACACACACACACACACACACACACACACACACACACACACACACACAC 1444

```

seq\_name: qb\_om:BTIE1A

seq\_documentation\_block:

LOCUS BTIE1A 4631 bp mRNA linear MAM 04-Nov-1993

DEFINITION B.taurus Tie 1 mRNA

ACCESSION X71423

VERSION X71423.1 GI:295575

KEYWORDS receptor tyrosine kinase.

SOURCE cow.

ORGANISM Bos taurus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Cetartiodactyla, Ruminantia, Pecora, Bovidae;

Bovidae; Bovinae; Bos.

1 (bases 1 to 4631)

Sato, T.N.

Direct Submission

Submitted (18-MAR-1993) T.N. Sato, Roche Institute of Molecular

Biology, 340 Kingsland St., Nutley, NJ 07035, USA

2 (bases 1 to 4631)

Sato, T.N., Ohtsuka, Y., Kozak, C.A., and Andus, K.L.

Tie-1 and Tie-2 define another class of putative receptor tyrosine

kinase genes expressed in early embryonic vasculature system

Proc Natl Acad Sci U S A 90 (26): 4335-4340 (1993)

94022374

Erratum: (published erratum appears in Proc Natl Acad Sci U S A

1993 Dec 15; 90(24):12056)

1993 Dec 15; 90(24):12056

Location: Val111-15

1, 4631

/organism="Bos taurus"

/db\_xref="taxon:9913"

/cell\_type="endothelial cells"

FEATURES

SOURCE

```

/tissue_type="brain"
/dev_stage="adult"
44..112
/gene="Tie 1"
44..3454
/gene="Tie 1"
/codon_start=1
/product="receptor tyrosine kinase"
/protein_id="CAA59554.1"
/db_xref="GI:296576"
/db_xref="SWISS-PROT:Q06805"
/translation="MWLEPPPLILPILPFLASHVCAAVDLTLLADLRLTETPFFETIV
SGEAGAGSGSIAMGPPILILPKDDRLVTPTRPMQPPHILARNGSRVTVRGSFSDILVG
VFSCVGGGTRVLYVHNSPGAILPKVTHVVKGDILAVLSARVKKQDILVWKSNG
SYEYLLRHEAQLGALLQLENLQVSSGILYSAIYLEASPLSSAFELKLVKCEARW
GQATKECPGLHNGVHDSGECVPEPFTGTGCEAHEHFGGSGUQPGTSGC
RLTELEPPYTSNGSSWSQSEACAPREGALHLQQLQNGSTGTRISGVCP
SCWHMCEKSDRLPQILDMVSELEFLNIMPRINCNAAGNPTVRGSMELKKPKDILV
LLSTKALVEPRTTARFVPRILALGDSGLWECHVSTSGGDSRRFRINVKVPIVPIIA
PRILAKSGRLVSPVVSFSDGPIASVRLHYRQDSTMAWSTLVVHSNVLMMNL
FKTGSVVALSPREGGAGWFFETMTTICPEPLKPLWLEWHVGGPRLVSWSL
PPVPGPLVGGFELLWDGARGGERENSVSPQATALLGLTPTGYGLDLYKHCT
LLGPASPAARVLLPPSGPPAPRILHQAALSDSEIQLMQRPENAAAPISKYIVEVVA
GSGPILMWIDVDRPEETSTIVRGINASTRYLFRVASVVOGPGHWSVFOSTLQNGLO
TECPQETHAAEGGLQGVIAVVGVSVALLLILAAALLLALIRKSCIHRRRTFYIO
SGSGETTLOFSQTLLTRPKPQPPPLNYPVLEWEDITFEDLIGEGNPGGVIRAMI
KKGLKNNAAIKMLKEYASNDHEDFAGLEVLCKLGHUINILNLGACENRGYLYIA
TEYAPYGNILDFLPKSPVETDPAPAFEPHASTISSPOLIPFASDAANMAYLSEPKO
FIURGLAARVVGWENLASKIACITGLSPFEVTVKTKMSEIYVHWMALESINVSVYTT
KSDVMSGVLLWEIVSGGIPYCOMTCAHLYERKLPQYRNGQPNRNDIKVYTHMKROCW
HDIYEREPFAQLAQIGRMLEARKAYVNMSTENFTYACIATAPEA"
44..3454
/gene="Tie 1"
113..3451
/gene="Tie 1"
/product="receptor tyrosine kinase"
/seq-09-733-764-2_copy_1_472 x BTIE1A
Align seq 1/1 to: BTIE1A from: 1 to: 3631
4 LeuAlaSerLeuValLeuCysGlyValSerLeuLeu..... 15
||| ||||| |||||
35 rProAspArgPheGlyCysLeuCysSerLeuCysSerLeuCysSerLeu 84
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
16 LeuSerGlyThrValGlyCysAlaMetAspLeuLeuLeuLeuLeuLeuLeu 42
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
85 CMTGGITTCALGILHSGSGSGSGTGGAGTGGAGTGTGGTGGAGGAG 144
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
32 euProLeuValSerAspAlaGluThrSerLeuThrCysIle..... 45
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
135 rProAspArgPheGlyCysLeuCysSerLeuCysSerLeuCysSerLeu 184
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
46 .....AlaSerGlyThrArgProHisGluProIleThr 56
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
185 rProAspArgPheGlyCysLeuCysSerLeuCysSerLeuCysSerLeu 228
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
56 rTieGlyArgAspPheGluAlaLeuMetAsnGlnHis.....GlnAsp 71
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
229 rProAspArgPheGlyCysLeuCysSerLeuCysSerLeuCysSerLeu 278
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
71 rLeuGluValThrGlnAspValThrArgGluTrpAlaLysLysValVal 87
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
279 CG .....CAACCGCGCGGCAAGCGGCGAGCGAGCGAGCGAGCGAGCGAG 314

```

alignment\_scores:

Quality: 874.00 Length: 474

Ratio: 2.766 Gaps: 13

Percent Similarity: 66.808 Percent Identity: 48.901

alignment\_block:

qs-09-733-764-2\_copy\_1\_472 x BTIE1A

Align seq 1/1 to: BTIE1A from: 1 to: 3631

```

88 TrpLysArg...GluLysAlaSerLysIleAsnGlyAlaIaTyrPheCysG1 103
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
314 GTCGGCGGCTTCTCCACAGCGCTCCACCTGGCGGCGGCTTCTCCGCGCT 363
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
103 wilyArqValArgGlyGluAlaIleArqIleArqThrMetLysMetArq 120
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
364 GAGC   GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 404
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
120 InGlnAlaSerPheLeuProAlaThrLeuThrMetThrValAspLysGly 136
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
405 CIGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 454
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
137 AspAsnValAsnIleSerPheLysLysValLeuIleLysGluGluAspAl 153
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
455 GACACGGCTGTACTTTCCGCGCGAGTCCGAGGAGAGAGAGAGAGAGAG 504
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
153 aValIleTyrLysAsnGlySerPheIleHisSerValProArgHisGluV 170
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
505 GATCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 554
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
170 alProAsp...IleGluGluValHisLeuProHisAlaGlnPro 183
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
555 CCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 598
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
184 GlnAspAlaGlyValTyrSerAlaArqTyrIleGlyGlyAsnLeuPheTh 200
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
599 TGTGTGAGCGGAGATGACAGTGGGACATACCTGGAGAGAGAGAGAGAG 648
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
200 rSerAlaPheThrArqLeuIleValArqArqCysGluAlaGlnLysTrpG 217
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
649 CAGTGGCTTCTTCACTCACTGAGAGAGAGAGAGAGAGAGAGAGAGAG 698
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
217 TyrGluCysAsnHisLeuCysThrAlaCysMetAsnGlyValCys 233
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
699 GACAAAGATGTACAAAGAAATAGAGAGAGAGAGAGAGAGAGAGAGAG 748
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
234 HisGluAspThrGlyGlyCysIleCysProGlyPheMetGlyArqTh 250
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
749 CAGAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 798
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
250 rCysGlyLysAlaCysGluLeuHisThrPheGlyArqThrCysLysGluA 267
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
799 CTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 848
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
267 rGlySerGlyGlnCysGlyCysLysSerTyrValPheCysLeuProAsp 283
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
849 AGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 898
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
284 ProTyrGlyCysSerCysAlaThrGlyTrpLysGlyLeuGlnCysAsnG1 300
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
899 CCTATGCGCTGCTCTGTGGATCTGGATGGAGAGAGAGAGAGAGAGAG 948
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
300 uAlaCysHisProGlyPheTyrGlyProAspCysLysLeuArqCysSerC 317
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
949 AGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 998
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
317 yAsnAsnGlnGluMetCysAspArqPheGlnGlyCysLeuCysSerPro 333
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
999 GTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1048
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
334 GlyTrpGlnGlyLeuHisCysGluArqGluGlyIleProArqMetThrPr 350
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1049 GGGTGGCAATGGGATGGACATGGACAGAGAGAGAGAGAGAGAGAGAG 1089
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
350 rLysIleValAspLeuProAspHisIleGluValAsnSerGlyLysPheA 367
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1090 CCAGATCTTGGACATGGTCTGAGAGCTGGAGTCAACTTAGACACAAATGC 1139
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
367 snProIle...CysLysAlaSerGlyTrpProLeuProThrAsnGlnGlu 382
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1140 CCGGATATAAATGTATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1189

```

```

383 MetThrLeuValLysProAspGlyThrVal...LeuHisProLysAspPh 398
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1190 ATCCAGCTTCGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1239
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
398 eAsnHisThrAspHisPheSerValAlaIlePheThrIleHisArgIleL 415
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1240 TGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1286
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
415 euProProAspSerGlyValTrpValCysSerValAsnThrValAlaGly 431
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1287 CTCTTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1336
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
432 MetValGlnLysProPheAsnIleSerValLysValLeuProLysProLe 448
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1337 CAGATATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1386
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
448 uAsnAlaProAsnValIle 454
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1387 GATTCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1405

```



OM 01: US 09-733-764-2\_COPY\_1\_472.rng N\_Geneseq\_032802 \* out\_format pfs  
 Date: Sep 23, 2002 11:25 PM  
 About: Results were produced by the Geneset software, version 4.5.  
 Copyright (c) 1993-2000 CompuGen Ltd.

## Command line parameters:

-MODEL=frame\_22n\_model -DEV=xlh  
 -O=/cdm2\_2/US09-733-764-2\_COPY\_1\_472.rng -FASTA=very\_fast\_1\_545  
 -DB=N\_Geneseq\_032802 -GFM=fastap -SUFFIX=ring -CAP=12-000  
 -GAPEXT=4-000 -MINMATCH=0-100 -L=0 -L=0 -L=0  
 -QCAP=4-500 -GAPEXT=0-050 -XCAP=10-000 -XGAPEXT=0-500  
 -FCAP=6-000 -DEPEXT=7-000 -STAPT=1 -MATRIX=DIOSUM62  
 -DELOP=6-000 -DEPEXT=7-000 -STAPT=1 -MATRIX=DIOSUM62  
 -TRANS=human40 rdi -T1T1=45 -LOCALIGN=200 -THP=SCORE-pct  
 -THK=MAX=100 -THK=MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFM=pfs  
 -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
 -USER=US09733764\_RCGNL\_1\_0 -NCPU=6 -ICPU=3 -LONGLOG  
 -DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -NO\_XLPXY -WAIT -THREADS=1

## Search information block:

Query: US-09-733-764-2\_COPY\_1\_472  
 Query length: 472  
 Database: N\_Geneseq\_032802:  
 Database sequences: 1736436  
 Database length: 85845721  
 Search time (sec): 221.93000

## score\_list:

Sequence	Strd Orig	ZScore	RScore	len	Documentation
/SIDSL/gcdata/geneseq/geneseq-emb1/NA1994.DAT:AA055179 +	2605.00	3317.35	2.0e-176	41	
/SIDSL/gcdata/geneseq/geneseq-emb1/NA1995.DAT:AA075334 +	2407.00	3064.64	2.4e-162	37	
/SIDSL/gcdata/geneseq/geneseq-emb1/NA1994.DAT:AA062129 +	2407.00	3063.68	2.7e-162	41	
/SIDSL/gcdata/geneseq/geneseq-emb1/NA2000.DAT:AA240648 +	2407.00	3063.68	2.7e-162	41	
/SIDSL/gcdata/geneseq/geneseq-emb1/NA2000.DAT:AA240640 +	2407.00	3063.67	2.7e-162	41	
/SIDSL/gcdata/geneseq/geneseq-emb1/NA1994.DAT:AA043740 +	2407.00	3063.27	2.4e-162	41	
/SIDSL/gcdata/geneseq/geneseq-emb1/NA1995.DAT:AA051999 +	2407.00	3062.71	2.0e-162	41	
/SIDSL/gcdata/geneseq/geneseq-emb1/NA1994.DAT:AA056647 +	2407.00	3045.74	2.7e-161	41	
/SIDSL/gcdata/geneseq/geneseq-emb1/NA1993.DAT:AA047064 +	865.50	1090.16	2.2e-52	38	
/SIDSL/gcdata/geneseq/geneseq-emb1/NA2001B.DAT:AA579371 +	688.50	859.02	1.7e-49	38	
/SIDSL/gcdata/geneseq/geneseq-emb1/NA1993.DAT:AA047065 +	675.50	847.14	7.7e-39	37	
/SIDSL/gcdata/geneseq/geneseq-emb1/NA2001B.DAT:AA119479 +	331.00	415.43	8.9e-15	11	
/SIDSL/gcdata/geneseq/geneseq-emb1/NA2001B.DAT:AA119878 +	331.00	406.48	2.7e-14	11	
/SIDSL/gcdata/geneseq/geneseq-emb1/NA2001B.DAT:AA104934 +	331.00	391.73	1.8e-13	11	
/SIDSL/gcdata/geneseq/geneseq-emb1/NA2000.DAT:AA241782 +	318.50	398.44	7.5e-14	14	
/SIDSL/gcdata/geneseq/geneseq-emb1/NA2001A.DAT:AA099715 +	318.50	398.44	7.5e-14	14	
/SIDSL/gcdata/geneseq/geneseq-emb1/NA2001A.DAT:AA099760 +	318.50	397.47	8.6e-14	14	
/SIDSL/gcdata/geneseq/geneseq-emb1/NA2000.DAT:AA261678 +	318.50	397.43	8.7e-14	16	
/SIDSL/gcdata/geneseq/geneseq-emb1/NA2001A.DAT:AA059461 +	318.50	397.43	8.7e-14	14	
/SIDSL/gcdata/geneseq/geneseq-emb1/NA2001A.DAT:AA02807 +	315.00	395.74	1.1e-13	14	
/SIDSL/gcdata/geneseq/geneseq-emb1/NA2001A.DAT:AA134844 +	315.00	391.24	1.9e-13	14	
/SIDSL/gcdata/geneseq/geneseq-emb1/NA2001B.DAT:AA568013 +	313.50	390.35	2.1e-13	14	
/SIDSL/gcdata/geneseq/geneseq-emb1/NA2001B.DAT:AA568013 +	313.50	385.11	4.2e-13	14	
/SIDSL/gcdata/geneseq/geneseq-emb1/NA2001B.DAT:AA572219 +	312.50	394.85	1.2e-13	14	
/SIDSL/gcdata/geneseq/geneseq-emb1/NA2001A.DAT:AA027790 +	309.00	386.11	3.7e-13	14	
/SIDSL/gcdata/geneseq/geneseq-emb1/NA2001A.DAT:AA027789 +	309.00	381.12	7.0e-13	14	
/SIDSL/gcdata/geneseq/geneseq-emb1/NA2001A.DAT:AA027810 +	309.00	380.67	7.4e-13	14	
/SIDSL/gcdata/geneseq/geneseq-emb1/NA2000.DAT:AA030650 +	308.00	385.09	4.2e-13	14	
/SIDSL/gcdata/geneseq/geneseq-emb1/NA2001A.DAT:AA127792 +	307.00	381.29	6.9e-13	14	
/SIDSL/gcdata/geneseq/geneseq-emb1/NA2001A.DAT:AA027791 +	307.00	375.55	1.4e-12	14	
/SIDSL/gcdata/geneseq/geneseq-emb1/NA2001A.DAT:AA027788 +	300.50	368.47	6.6e-12	14	
/SIDSL/gcdata/geneseq/geneseq-emb1/NA2001A.DAT:AA027787 +	300.50	364.07	6.2e-12	14	
/SIDSL/gcdata/geneseq/geneseq-emb1/NA2001A.DAT:AA020811 +	300.00	373.45	1.9e-12	14	
/SIDSL/gcdata/geneseq/geneseq-emb1/NA2001B.DAT:AA108119 +	299.00	369.16	3.2e-12	14	
/SIDSL/gcdata/geneseq/geneseq-emb1/NA2000.DAT:AA051078 +	298.00	352.30	2.9e-11	14	
/SIDSL/gcdata/geneseq/geneseq-emb1/NA2001A.DAT:AA027469 +	297.50	351.46	4.2e-11	14	
/SIDSL/gcdata/geneseq/geneseq-emb1/NA2001B.DAT:AA051078 +	294.50	357.73	1.4e-11	14	
/SIDSL/gcdata/geneseq/geneseq-emb1/NA2001B.DAT:AA027362 +	291.50	354.07	2.2e-11	14	
/SIDSL/gcdata/geneseq/geneseq-emb1/NA2001B.DAT:AA086742 +	277.00	357.41	1.5e-11	14	

/SIDSL/gcdata/geneseq/geneseq-emb1/NA2001B.DAT:AA051078 + 277.00 357.41 1.5e-11  
 /SIDSL/gcdata/geneseq/geneseq-emb1/NA2002.DAT:AA045098 + 277.00 340.04 1.4e-10  
 /SIDSL/gcdata/geneseq/geneseq-emb1/NA2001A.DAT:AA027488 + 277.00 337.43 1.9e-10  
 /SIDSL/gcdata/geneseq/geneseq-emb1/NA2001A.DAT:AA027486 + 277.00 447.40 1.9e-10  
 /SIDSL/gcdata/geneseq/geneseq-emb1/NA2001B.DAT:AA086746 + 272.50 339.92 1.4e-10

seq\_name: /SIDSL/gcdata/geneseq/geneseq-emb1/NA1994.DAT:AA055179

seq\_documentation\_block:

US AA055179 Standard: cDNA to mRNA; 4138 BP.  
 AC AA055179;  
 ET 25-JUN-1994 (first entry)  
 DE Human orphan receptor kinase gene.  
 KW Ork gene, ligands, antibodies, PCR, amplification; ss.  
 OS Homo sapiens.

Feature	Key	Location/Qualifiers
CD5	149..3523	
misc_feature	/*tag= a	2744..3219
misc_feature	/*tag= b	/*note= "amplified by PCR"
misc_feature	/*tag= c	149..3520
misc_feature	/*note= "specifically claimed fragment"	203..3520
misc_feature	/*tag= d	149..2383
misc_feature	/*note= "specifically claimed fragment"	/*tag= e
misc_feature	/*note= "specifically claimed fragment"	203..2383
misc_feature	/*tag= f	/*note= "specifically claimed fragment"

W09400469-A.  
 06-JAN-1994.  
 25-JUN-1993; 94W0-0506094.  
 26-JUN-1992; 92US-0905600.  
 (IMMUNEX CORP.)  
 Ziegler SF;

WT: 1994-026132/03.  
 F-PSDB: AAP45440.  
 DNA and protein sequences for orphan receptor tyrosine kinase -  
 and expression vectors for produ. of recombinant protein and  
 antibodies specific for the protein, useful in research  
 Claim 1, Fig 1, 57pp; English.

Generate oligonucleotide primers based on the sequence disclosed  
 in the kinase domain of all receptor tyrosine kinases was used for  
 PCR of single stranded cDNA from human placental polyA mRNA. PCR  
 prod. HKK-6 contained a novel sequence which was used as a probe to  
 isolate longer fragments from a human placental cDNA library. One  
 clone (sequence shown) contained the entire coding region and was  
 called the ork gene. The ork products can be used as a search tool  
 in in vitro assays for detection of ork, its ligands or their  
 interactions.

Sequence 4138 BP, 1170 A, 911 C, 987 G, 1070 T, 0 other;

alignment\_scores:  
 Quality: 2605.00 Length: 472  
 Ratio: 5.519 Gaps: 0  
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:  
 us-09-733-764-2\_copy\_1\_472 x AAQ55179 ..  
 Align seq 1/1 to AAQ55179 from: 1 to: 4148

```

1 MetAspSerLeuAlaSerLeuValLeuGlyValSerLeuLeuLeuLeu 17
149 ATGGACCTTTTAACTTAACTTAACTTAACTTAACTTAACTTAACTT 198
17 rGlyThrVal.GluGlyAlaMetAspLeuLeuLeuLeuLeuLeuLeu 34
199 TGGACCTTGGAGGTGGATGATGATGATGATGATGATGATGATGATG 248
34 euValSerAspAlaGluThrSerLeuThrCysIleAlaSerGlyTrpArg 50
249 TTGTATCTCATCTGCAAACTCTCAGCTGCACTGCACTGCACTGCGC 298
51 ProHisGluProIleThrIleGlyArgAspPheGluAlaLeuMetAsnG 67
299 GCGCATGAGGACATCAGCATAGAGAGAGACCTTGAAGGCTTAATGAAC 348
67 nHisGluAspProLeuGluValThrGluAspValThrArgGluTrpAla 84
349 GCAACAGATCAAGCTGGAAGTAACTCAAGATGTCAGCAGAGAAATGG 398
84 ySLysValValTrpLysArgGluLysAlaSerLysIleAsnGlyAlaTyr 100
399 AAAAAGCTCTTTTGGAGACACAGAAAAGCTAGTAAATCAATTCGCTT 448
101 PheCysGluGlyArgValArgGlyGluAlaIleArgIleArgThrMetL 117
449 TTCTTTGAAAGGGGAGTTTCAGGAGAGAGATATAGGATATGAAACAT 498
117 sMetArgGlnGlnAlaSerPheLeuProAlaThrLeuThrMetThrVal 134
499 GATGGTTCATACAGTTCCTTCTCTATATATATATATATATATATAT 548
134 sPLysGlyAspAsnValAsnIleSerPheLysValLeuLeuLysGlu 150
549 ACAAGGACATACCGTACATATCTTCAAAAAGCTATTCATTAATTAAT 598
151 GluAspAlaValIleTyrIleAsnGlySerPheIleHisSerValProAr 167
599 GAAGATGCACTGATTAACTTAACTTAACTTAACTTAACTTAACTTAA 648
167 qHisGluValProAspIleLeuGluValHisLeuProHisAlaGlnPro 184
649 GATCAAGCTAGCTGATATCTTGAAGAGATATATGAGCTATATATAG 698
184 InAspAlaGlyValTyrSerAlaArgTyrIleGlyGlyAsnLeuPheThr 200
699 AGGATGCTGGACCTGATCTGCGGACAGTATATAGAGGAAACCTCTTCA 748
201 SerAlaPheThrArgLeuIleValArgArgCysGluAlaGlnLysTrpGl 217
749 TCGGGCTTAACTAGATATATATATATATATATATATATATATATAT 798
217 PheCysGlyAsnHisLeuGlyThrAlaCysMetAspAsnGlyValCysH 234
799 ACCTCAATGTAACATATCTTATATATATATATATATATATATATAT 848
234 IsGluAspThrGlyGlyCysGlyCysProGlyPheMetGlyArgThr 250
849 ATCAACATATCGACAAATATATATATATATATATATATATATATAT 898
251 CysGluLysAlaCysGluLeuHisThrPheGlyArgThrCysLysGluAr 267

```

```

899 TGTGAGAGAGGTTTGTCAACTGCACACAGCTTTGCAGAACTTTGTAAGA 948
267 gCysSerGlyGlnGluGlyCysLysSerTyrValPheCysAlaIleAsp 284
949 GTGACGTGCAATAAAGCAATGTAAGTCTTAATGCTGTGTGCTGCTGCT 998
284 rCysGlyCysSerCysAlaThrGlyTrpLysGlyLeuGlnCysAsnGlu 300
999 CCTATGGGCTTCTCTGTCGACACAGCTGCAAGGCTGCTGCTGCTGCT 1048
301 AlaCysHisPheGlyPheTyrGlyPheAspCysLysLeuArgCysSerCy 317
1049 GATATGCAATCTGCTTTTATAGGCAATATGTAAGTTATATATATAT 1098
317 sAsnAsnGlyGluMetCysAspArgPheGlnGlyCysLeuCysSerProG 334
1099 CAACATGGGAGACATGTCGATCGCTTCCAAAGGATGCTCTGCTCTCC 1148
334 TyrPheGlnGlyLeuGlnCysLeuArgGlnGlyIlePheArgMetThrPro 350
1149 GATGTAAGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1198
351 LysIleValAspLeuProAspHisIleGluValAsnSerGlyLysPheAs 367
1199 AAGATAGTGGATTTCCTATATATATATATATATATATATATATATAT 1248
367 nProIleCysLysAlaSerGlyTrpProLeuProThrAsnGluMetC 384
1249 TCTTATCTTAAAGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1298
384 hrLeuValLysProAspGlyThrValLeuHisProLysAspPheAsnHis 400
1299 CCTGCTGTAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1348
401 ThrAspHisPheSerValAlaIlePheThrIleHisArgIleLeuProP 417
1349 AGGATCAATTTCTATATATATATATATATATATATATATATATAT 1398
417 oAspSerGlyValTrpValCysSerValAsnThrValAlaGlyMetValG 434
1399 TGACCTCAGGAGTTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1448
434 LuLysProPheAsnIleSerValLysValLeuProLysProLeuAsnAla 450
1449 AAAAGCCCTTCAACATTTCTGCTTAAAGTCTGCTGCTGCTGCTGCT 1498
451 ProAsnValIleAspThrGlyHisAsnPheAlaValIleAsnIleSerSe 467
1499 CCMAACGCTGATTGACACTGGACATAAATCTTCTGCTCATCAACATCA 1548
467 rGluProLysPheGly 472
1549 TGAGCTTACTTTGGG 1564

```

seq\_name: /SIDSL/gcdata/geneseq/geneseq-emb1/NA1995.DAT:AAQ75334

seq\_documentation\_block:

XX AAQ75334 standard; cDNA; 3760 BP.

XX AAQ75334;

XX AC

XX XX

XX 24-Apr-1995 (first entry)

XX Marine tyrosine kinase receptor tie-2 cDNA.

XX Marine tyrosine kinase receptor; tie-2; hemolysis diaphosis; ds.

XX Mus musculus.

XX Key: Location/Qualifiers

XX CDS 152..3523

XX FT seq\_peptide 152..211

```

FT mat_peptide /*tag= b
FT 212...3520 /*tag= c
PN JP06315382-A.
XX 15-NOV-1994.
XX 06-MAY-1993: 933P-0129912.
XX 06-MAY-1993: 933P-0129912.
XX (SUDA/) SUDA IT.
XX (YAMA ) YAMAGUCHI PHARM CO LTD.
XX WP1: 1995-032331/05.
XX P-PSDB: AAR67391.
XX DNA coding a tie-2 receptor and a tie 2 receptor used in the
XX diagnosing hematosis
XX Claim 2, Pages 13-17, 17pp, Japanese.
XX AA075334 encodes AAR67391 the murine tyrosine kinase receptor tie-2.
XX The cDNA is thought to participate in hematosis, and can therefore
XX be used in the diagnosis of hematosis.
XX Sequence 3760 BP: 1024 A: 876 C: 970 G: 890 T: 0 other:
XX
XX alignment_scores:
XX Quality: 2407.00 Length: 472
XX Ratio: 5.233 Gaps: 0
XX Percent Similarity: 97.458 Percent Identity: 91.949
XX
XX alignment_block:
XX US-09-733-764-2_copy_1_472 x AA075334
XX
XX Align seq 1/1 to: AA075334 from: 1 to: 3760
XX
XX 1 MetAspSerLeuAlaSerLeuValLeuCysGlyValSerLeuLeuLeu 17
XX 152 ATGGACTCTTTAGCGGCTTAGTTCCTGCGGAGGACGCTTCCCTTCA 201
XX 17 rGlyThrValGlyGlyAlaMetAspLeuLeuLeuLeuLeuLeuProL 34
XX 202 TGAGATAGTAAAGGTGGATGAGAGTATGATTTGATTAATTCCTACCTC 251
XX 34 euValSerAspAlaGluThrSerLeuThrCysIleAlaSerGlyTrpArg 50
XX 252 TTGTGTCTGATGAGGAGAAAGATCTCAATGATTTGCTGCTGGTAC 301
XX 51 ProHisGluProIleThrIleGlyArgAspPheGluAlaLeuMetAsnG 67
XX 302 CCCCATGAGCCATCAGCATAGGAAGGACCTTTGAAGCCCTTAATGAACCA 351
XX 67 mHisGluAspProLeuGluValThrGlnAspValThrArgGluTrpAlaL 84
XX 352 GCACCAAGATCCTAGGAGGTATCTAAAGATGTAAAGAAATGGGCGA 401
XX 84 ystysValValTrpIysArgGluIysAlaSerIlystIleAsnGlyAlaTr 100
XX 402 AAAAAATTGTGGAGAGAGAGAAAGCCAGTAGATTAAATGGTCTAT 451
XX 101 PheCysGluGlyArgValArgGlyGluAlaIleArqIleArgThrMetLy 117
XX 452 TTCTGTGAAGCTCCAGTTCCAGGACAGCTATAAGCATACCGACCATGAA 501
XX 117 sMetArgGlnGlnAlaSerPheLeuProAlaThrLeuThrMetThrVal 134
XX 502 GATGGGTAAAGAGGTCTTCTTCAATGATTAATTTAATCATGACAGTGG 551
XX 134 splysGlyAspAsnValAsnIleSerPheIysIysValLeuIleCysGlu 150

```

```

552 ACAGGATCAATAAAGTCAAAATAATCTTCAAAAAAGCTGTTAATAAGAA 601
151 GluAspAlaValIleTyrIlysAsnGlySerPheIleHisSerValProAr 167
602 GAAGATGCAAGTATTTATATAAAAGTGAATCTTCAAGATCAAGAGGAGG 651
167 gHisGluValIProAspIleLeuGluValHisLeuProHisAlaGlnProG 184
652 GATCAAGTACCTGATATTTAGAAATCTACTTGAAGATGCTCAAGGAC 701
184 InAspAlaGlyValIYrSerAlaArgfyrIleGlyGlyAsnLeuPheThr 200
702 AGCAATGCTGCTGCTAGTGGGTCAGGTACATAGGAGCAAAACCTGTTCAG 751
201 SerAlaPheThrArgLeuIleValArqCysGluAlaGlnLysTrpG 217
752 TCAGCTCTTCACTAGGTTCATGTTGAGAAATGAGAAATACAAAGAGGAG 801
217 yProGluCysAsnHisLeuCysThrAlaCysMetAsnAspGlyValCysH 234
802 GAGAAATCTAGAGGCTCTTGTACTATTTGAAATATGAGAGTCTGAGC 851
234 isCluAspThrGlyGluCysIleCysIleCysProGlyPheMetGlyArgThr 250
852 ATCAAGATACCGGGGAAATGCAATTCGCCCTCCCTGGGCTTATATGGGAGCA 901
251 CysGluLysAlaCysGluLeuLeuHisThrPheGlyArgThrCysLysGluAr 267
902 TGTGAGAAAGCTTGTGAGCCGCACACATTTGGCAGGACCTGTAAAGAAAG 951
267 gCysSerGlyGlnGluClyCysLysSerIyrValIleCysLeuProAspF 284
952 GTGTAGTACAGAGAGAGATGATGATGCTTCAAGGATGCTCTCTCTCAAG 1001
284 rGlyTrpCysSerCysAlaThrGlyTrpIysGlyLeuGlnCysAsnGlu 300
1002 CTTCAGGGGTCTTCTGCTGCTTCAAGTGGAGGAGGAGTGGAGTGAAGAA 1051
301 AlaCysHisProGlyPheTyrGlyProAspCysLysLeuArgCysSerCy 317
1052 GCATGCTCATCTGCTTACTACGACACAGACTGTAACTCAGCTGCCACATG 1101
317 sAsnAsnGlyGluMetCysAspArgPheGlnGlyCysLeuCysSerProG 334
1102 TACCAATGAAGAGATATGATGCTGCTTCAAGGATGCTCTCTCTCAAG 1151
334 IYTrpGlnGlyLeuGlnCysGluArgGluGlyIleProArgMetIhrPro 350
1152 GATGGCAAGGGCTGACAGTGTGAGAAAGAGAGCCAGCCCAAGGATGACTCA 1201
351 LysIleValAspLeuProAspHisIleGluValAsnSerGlyLysPheAs 367
1202 CATATAGAGGATCTGACAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 1251
367 nProIleCysLysAlaSerGlyTrpProLeuProThrAsnGluGluMetI 384
1252 CCGCATCTGTAAAGGCTCTGGGTGGTCACTACCTACTAGTGAAGAAATGA 1301
384 hLeuValIysProAspGlyThrValLeuHisProIysAspPheAsnHis 400
1302 CCTACTGAAGCCAGATGGGACAGTCTCCCAACCAAAATCATTCAACTAT 1351
401 ThrAspHisPheSerValAlaIlePheThrIleHisArqIleLeuProPr 417
1352 ATCAATGCTTCTCAGTGGGCTATATGCTGCTCAACCTGCTTACCTGCC 1401
417 oAspSerGlyValITrpValICysSerValAsnThrValAlaGlyMetVal 434
1402 TGAATCAAGAGTCTGAGTCTTGAATGTAAGTGAATGTAAGTGAATGTA 1451
434 IulysProPheAsnIleSerValLysValLeuProIysProLeuAsnAla 450

```









```

324 GCACCAATATCTATTCAGAGGATTAATTAATATCTAATAGAGAAATBGGAGA 373
84 yslsValValTrpIysArgGlyAlaSerGly::IcAsnGlyAlaIyr 100
|||||
374 AAAAAGTCTTTGCAAGAGAGAAAAGCCAGTAAGATTAATGCTGCTAT 423
101 PheCysGluGlyArgValArgGlyGluAlaIleArqIleArgThrMelly 117
|||||
424 TTCCTGAAGTCCAGTTCGAGGACAGGCTATAAGGATACGACCATGAA 473
117 sMetArgGlnGlnAlaSerPheLeuProAlaThrLeuThrMetThrValA 134
|||||
474 GATGGGTCAAATATCTTCTATCTATCTATCTATCTATCTATCTATCTAT 523
134 splysGlyAspAsnValAsnIleSerPheIysValIleuIleIysGlu 150
|||||
524 ACAGGAGAGATAATCTCAACATATCTTCAAAAAGGCTGTTAAATTAAGAA 573
151 GluAspAlaValIleTyrLysAsnGlySerPheIleIleIleSerValProAr 167
|||||
574 CAAGATCGAGTCATTTACAAAATGGCTCCTTCATCCACACAGTGCCTCCG 623
167 gHisGluValProAspIleLeuGluValHisLeuProHisAlaGlnProG 184
|||||
624 GATCTAAATATCTATATTTAAATTTAAATTTAAATTTAAATTTAAATTT 673
184 InAspAlaGlyValIleTyrSerAlaArgIleTyrIleGlyGlyAsnLeuPheThr 200
|||||
674 AGGATGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 723
201 SerAlaPheThrAspLeuLeuValArgGlyCysGluAlaIleLysIleP 217
|||||
724 TCAGCTTCACACAGGTCATCTTCGAGATGCGAAGCTCAGAACTGGGG 773
217 yProGluCysAsnHisLeuCysThrAlaCysMetAsnAsnGlyValCysH 234
|||||
774 GCTGACGAGTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 823
234 IsGlnAspThrGlyGlyCysGlyCysThrAlaPheMetGlyArgThr 250
|||||
824 ATGAAGATACCGGGAATGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCT 873
251 CysIleLysAlaCysGluLeuHisThrPheGlyArgThrCysGlyArg 267
|||||
874 TCTCAGAAAGCTTGAGACCGCAACATTCGACGAGGCTGCTGCTGCTGCT 923
267 gCysSerGlyGlnGluGlyCysLysSerTyrValPheCysLeuProAsp 284
|||||
924 GTGTATGTAATAGAAATATATATATATATATATATATATATATATAT 973
284 rOTyrGlyCysSerCysAlaThrGlyTrpIysGlyLeuGlnCysAsnGlu 300
|||||
974 CTTACGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1023
301 AlaCysHisProGlyPheTyrGlyProAspCysLysLeuArgCysSerCy 317
|||||
1024 CATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1073
317 sAsnAsnGlyGluMetCysAspArgPheGlnGlyCysLeuCysSerProG 334
|||||
1074 TACCAATTAATAGAAATATATATATATATATATATATATATATATAT 1123
334 TyrTrpGlnGlnCysGluArgGlyIleThrAlaGluMetThrPro 350
|||||
1124 CATGCAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1173
351 LysIleValAspLeuProAspHisIleGlyValAsnSerGlyLysPheAs 367
|||||
1174 CAGATAGAGGATTTGCGACATCATCTGAAGTAAACAGTGGAAAAATTTAA 1223
367 nProIleCysLysAlaSerGlyTyrTrpLeuProThrAsnGluGluMet 384
|||||
1224 GCGCATCTG:AAAAGTCTGGAGTGGTAAATATATAGTAAAGAAATGA 1273

```

```

384 hrLeuValLysProAspGlyThrValLeuHisProLysAspPheAsnHis 400
|||||
1274 CCTACTGAAGCCAGATGGACAGTCTCCCAACCAATCATTCAACTAT 1323
401 ThrAspHisPheSerValAlaIlePheThrIleHisArgIleLeuProPr 417
|||||
1324 ACAGATGTTTCTCAATGTCATATATATATATATATATATATATAT 1373
417 oAspSerGlyValTrpValCysSerValAsnThrValAlaGlyMetValG 434
|||||
1374 TGAATCAGAAATCTGATCTGATCTGATCTGATCTGATCTGATCTG 1423
434 LuLysProPheAsnIleSerValLysValLeuProLysProLeuAsnAla 450
|||||
1424 AAAAGCTTTTCAACATTTCTGTCACAAAGTTCTTCCAGAGGCTGCG 1473
451 ProAsnValIleAspThrGlyHisAsnPheAlaValIleAsnIleSerSe 467
|||||
1474 CTAAATGTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1523
467 rGluProTyrPheGly 472
|||||
1524 TGAGCTTACCTTGGG 1539
seq_name: /SIN1/gcdata/geneseq/geneseq-emb1/NA1694 DAT-AA1694
seq_documentation_block:
ID AA063740 standard; cDNA; 4364 BP.
XX
AC AA063740;
XX
DI 17-OCT-1994 (first entry)
XX
DE Protein tyrosine kinase JIL.
XX
KW Protein tyrosine kinase, PTK, receptor tyrosine kinase, PTK;
KW JIL; carcinoma; cancer therapy; fibronectin type III; FNIII; SS.
XX
OS Mus sp.
XX
FH Key Location/Qualifiers
FT CDS 16..3384
FT /*tag- a
XX
PW W09410197-A.
XX
PD 11-MAY-1994.
XX
PF 29-OCT-1994; G3WO-A000560.
XX
PR 30-OCT-1992; G2AU-0005581.
XX
PA (LUDW-) LUDWIG INST CANCER RES.
XX
PI Runling AS, Wilks AF;
XX
DR WPI; 1994-167386/20.
XX
PT New protein tyrosine kinase, corresp DNA and ligands - useful as
PT anticancer agents in treatment of carcinoma
XX
PS disclosure, fig 1, 60pp, English.
XX
CC Poly A+ mRNA from neonatal mouse colon was used to generate cDNA
CC that was amplified by PCR. Clones were sequenced, and the pJIL
CC clone was used to screen a mouse lung cDNA library. Overlapping
CC clones were sequenced to generate the sequence given in AA063740.
CC JIL clone encodes a new protein tyrosine kinase that includes an
CC extracellular domain comprising at least 3 fibronectin type III
CC repeats, an immunoglobulin-like domain and an epidermal growth
CC factor-like domain.
XX

```

The

SO Sequence 4364 bp: 1267 A; 965 G; 1066 C; 1099 T; 2 other.

# alignment\_scores:

Quality: 2407/00 Length: 472  
Ratio: 5.243 Gaps: 0  
Percent Similarity: 97.458 Percent Identity: 91.949

## alignment\_block:

us-09-733-764-2\_copy\_1\_472 x AA063730  
Align seq 1/1 to: AA063730 from: 1 to: 4364  
1 MetAspSerLeuAlaSerLeuValLeuCysGlyValSerLeuLeuLeu 17  
16 ATGGACTCTTACGGGAGTTATGTCGTGGAGTCACCTTGGCTCTTAA 65  
17 rGlyThrValGluGlyAlaMetAspLeuLeuLeuLeuLeuLeu 34  
66 TGCAGTACAGAAAGGTGGCAAGCACTGATCTGATCAATTCGCTAGCTC 115  
34 euValSerAspAlaGluThrSerLeuThrCysIleAlaSerGlyTrpArg 50  
116 TTGGTGCTGATGCTAAACATGCTTACCTGCTGATTCGCTGCGGAC 165  
51 ProHisGluProIleThrIleGlyArgAspPheGluAlaLeuMetAsnGI 67  
166 GCGCATGAGCGCATACCATATAGAGAGACTTTGAAAGCTTAATGAAACA 215  
67 nHisGlnAspProLeuGluValThrGlnAspValThrArgGluTrpAlaL 84  
216 GCACCAAGATCCACTGGAGGTACTCAAGATGTGACCAGAAATGGGCGA 265  
84 yLysValValIleTrpLysArgGluLysAlaSerLysIleAsnGlyAlaTr 100  
266 AAAAAAGTCTTGGAGAGACAGAAAGGCCAGTAAGATTAATGCTGCTTAT 315  
101 PheCysGluGlyArgValArgGlyGluAlaIleArgIleArgThrMetLy 117  
316 TCTGTGAAGAGTGAGATTTGAGAGAGATATAGATATAGATATAGAG 365  
117 sMetArgGlnGlnAlaSerPheLeuProAlaThrLeuThrMetThrValA 134  
366 GATCCCTCAACAAGCGCTCTCTTACCCTGCTACTTTAACTAAGACCGTGG 415  
134 sPylsGlyAspAsnValAsnIleSerPheLysLysValLeuIleLysGlu 150  
416 ACAGGGGAGATTAATGTGAACAATATCTTCAAAAAGGCTGTAACTAAAG 465  
151 GluAspAlaValIleTrpLysAsnGlySerPheIleHisSerValTrpArg 167  
466 GAAATATGATGATTTTAAATAAGATGCTTCTTCAATATAGAGAGAG 515  
167 gHisGluValProAspIleLeuGluValHisLeuProHisAlaGlnProG 184  
516 GATCAAGAGTATGATATACAACTTCACTTCACTTCACTTCACTTCA 565  
184 IaAspAlaGlyValTrpSerAlaLysTrpIleGlyGlyAsnLeuPheThr 200  
566 AGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 615  
201 SerAlaPheThrArgLeuIleValArgArgCysGluAlaGlnIleTrpGI 217  
616 TCAATTTTAAATGATATGATGATGATGATGATGATGATGATGATG 665  
217 gProGluCysAsnHisLeuGlyThrAlaCysMetAsnGlyValCysH 234  
666 GCGCGACTGTAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 715  
234 iGluAspThrGlyGluCysIleCysProProGlyPheMetGlyArgThr 250  
716 ATGAGATATAGCGGGAATGCAATTCGCTGCTGCTGCTGCTGCTGCTG 765

251 CysGluLysAlaCysGluLeuLeuHisThrPheGlyArgThrCysLysGluArg 267  
766 TGTGATAAATATTGACAGAGAGAGATATTTTGGAGAGACCTGTAAAGAAAG 815  
267 gCysSerGlyGlnGluGlyCysLysSerTyrrValPheCysLeuProAspF 284  
816 GTGTACTGGACATCAAGAGATGCAAGTCTTATGTGTGTGTGTGTGTGT 865  
284 rGlyGlyCysSerCysAlaThrGlyTrpLysGlyLeuGlnCysAsnGlu 400  
866 TTATAGAGGTCTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 915  
301 AlaCysHisProClyPheClyTrpGlyProAspCysLysCysArgCysSerC 317  
916 GTATATGCTATCTGATTAATAGAGAGAGAGATGTAAGATCAAGTGTGCTG 965  
317 sAsnAsnGlyGluMetCysAspArgPheGlnGlyCysLeuCysSerProG 334  
966 TACCATATGACAGATATGATGATGATGATGATGATGATGATGATGATG 1015  
334 lYrTrpGlnGlyLeuGlnCysGluArgGluGlyIleProArgMetThrPro 350  
1016 GATGATAGAGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1065  
351 LysIleValAspLeuProAspHisIleGluValAsnSerGlyLysPheAs 367  
1066 CATATAGAGATTTGCTGATCAATATGAAATGAAATGAAATGAAATTTAA 1115  
367 nProIleCysLysAlaSerGlyTrpProLeuProThrAsnGluGluMet 384  
1116 CCGCATATGCAAAAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1165  
384 hrLeuValLysProAspGlyThrValLeuHisProLysAspPheAsnHis 400  
1166 CCGTACTGAGAGATGATGAGAGATGAGAGATGAGAGATGAGAGATGAG 1215  
401 ThrAspHisPheSerValAlaIlePheThrIleHisArgIleLeuProPr 417  
1216 ACAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1265  
417 oAspSerGlyValTrpValCysSerValAsnThrValAlaLysMetValG 434  
1266 TCATCTAGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1315  
434 LuLysProPheAsnIleSerValLysValLeuProLysProLeuAsnAla 450  
1316 AAAAGCTTTTAAATTTTCTCAAGTCTTCTCAAGTCTTCTCAAGTCTTCT 1465  
451 ProAsnValIleCysPheThrGlyHisAsnPheAlaValIleAsnIleSer 467  
1366 CCAATATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1415  
467 rGluProTrpPheGly 472  
1416 TCAGCGCTTACTTGGG 1431

seq\_name: /SISU/qcdata/geneseq/geneseq-emb1/NA1995.DAT:AA091999

seq\_documentation\_block:

ID AA091999 standard; DNA; 4640 bp.

XX AA091999;

XX 21-JAN-1996 (first entry)

XX Mouse tic 2 receptor tyrosine kinase DNA.

XX tic-2; receptor tyrosine kinase; DNA primer; cancer; angiogenesis;

KW vasculogenesis; SS.

XX Mus musculus.

XX Key

XX Location/Qualifiers

```

FT CDS 341..3712
PT /*tag= a
XX W09513387-A1.
XX 18-MAY-1995.
XX 12-NOV-1994: 94W0-EP03767.
XX 12-NOV-1993: 930S-0152552.
XX (PLAC ) MAX PLANCK GFS FÖRDERUNG WISSENSCHAFTEN.
XX Risau W.
XX WPI: 1995-194105/25.
XX P-PSDB. AAR73951.
XX New tie-2 receptor tyrosine kinase and related nucleic acid - and
XX methods for detecting tie-2 modulators for treating eg cancer,
XX associated with angiogenesis and vasculogenesis
XX Claim 3, Page 42, 81pp, English.
XX this DNA may be expressed recombinantly in a host cell. The DNA was
XX isolated from mouse brain and can be used as a probe to detect
XX related genes of other animals or expressed in host cells e.g. E
XX coli. The expressed protein may be used to treat diseases, or
XX processes, associated with angiogenesis and vasculogenesis, or
XX cancer. Cells that express the protein are used in screening
XX procedures and recombinant protein can be used for affinity
XX purification of tie-2 ligand. The DNA may be used diagnostically to
XX detect gene expression.
XX Sequence 4640 BP: 1403 A: 1041 C: 1147 G: 1149 T: 0 other:
XX
XX alignment_scores:
XX Quality: 2407.00 Length: 472
XX Ratio: 5.233 Gaps: 0
XX Percent Similarity: 97.458 Percent Identity: 91.949
XX
XX alignment_block:
XX US-09-733-764-2_copy_1_472 x AAG91999 ..
XX Align seq 1/1 to: AAG91999 from: 1 to: 4640
XX
XX 1 MetAspSerLeuAlaSerLeuValLeuCysGlyValSerLeuLeuLeuLeu 17
XX |||||
XX 341 ATGGACTCTTTAGCGGGCTTACTTCTCTGTCGAGTCACGCTTCCTCTTTA 390
XX
XX 17 cGlyThrValGluGlyAlaMetAspLeuLeuLeuLeuLeuLeuLeuLeu 34
XX |||||
XX 391 TGGAGTAGTAGAGGCGGACCGACCGGATCTTGATCAATTCCTCCTACCTC 440
XX
XX 34 euValSerAspAlaGluThrSerLeuThrCysTleAlaSerGlyTyrArg 50
XX |||||
XX 441 TTGTGTCTATGAGGAGAAATATCTCTCTCTCTCTCTCTCTCTCTCTCT 490
XX
XX 53 ProHisGluProTleThrTleClyArgAspPheGluAlaGluMetAsnCl 67
XX |||||
XX 491 GCGATGAGCGGCGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 540
XX
XX 67 HisGlnAspProLeuGluValThrGlnAspValThrArgGluTrpAlaL 84
XX |||||
XX 541 GATTAAGAGATCTATGAGGATCTATGAGGATCTATGAGGATCTATGAG 590
XX
XX 84 yLysValValTrpLysArgGluLysAlaSerLysTleAsnGlyAlaTyr 100
XX |||||
XX 591 AAAAAGTTCTTTGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 640
XX
XX 101 PheCysGluGlyArgValArgGlyGluAlaAlaCargileArgThrMetIy 117
XX |||||

```

```

641 TTCGTGAAGGCTGAGCTGAGGAGACAGGCTATAGGATACAGGACCA 690
117 sMetArgGlnGlnAlaSerPheLeuProAlaThrLeuThrMetThrValA 134
|||||
691 GATGCTCTAAAGAGGCTTCTTAACTGTAATTAATTAAGGAGGAG 740
134 sPheGlyAspAsnValAsnTleSerPheLysLysValLeuLysGlu 150
|||||
741 ACAGGGAGATATGTCAGCAATATCTTCAAAAGGAGTAAATTAAGAA 790
151 GluAspAlaValIleTyrLysAsnGlySerPheIleHisSerValProAr 167
|||||
791 CAAGATCGAGTGATTACAAAATGGCTCTTCATCCACTCAGTGGCCG 840
167 qHisGluValProAspIleLeuGluValHisLeuProHisAlaGlnProG 184
|||||
841 GCAICAAATATATATATTTAAAGATCAATGAGGAGATGAGGAGG 890
184 tAspAlaGlyValTyrSerAlaGlyTyrIleGlyCysAsnGluLeu 200
|||||
891 AATATATGAGGAGTATATGAGGAGTATATGAGGAGTATATGAGGAG 940
201 SerAlaPheThrArgLeuIleValArgCysGluAlaGlnLysTrpGl 217
|||||
941 TCAGGCTTCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 990
217 yProGlyCysAsnHisLeuCysThrAlaCysMetAsnAsnLysValCysH 234
|||||
991 GAGGAGTATATGAGGAGTATATGAGGAGTATATGAGGAGTATATGAG 1040
234 tAspAlaGlyValTyrSerAlaGlyTyrIleGlyCysAsnGluLeu 250
|||||
1041 ATGAGGATATGAGGAGTATATGAGGAGTATATGAGGAGTATATGAGGAG 1090
251 CysGluLysAlaCysGluLeuHisThrPheGlyArgThrCysLysGlnAr 267
|||||
1091 TGTGAGAAAGCTTCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1140
267 qCysSerGlyValLeuLysLysSerTyrValPheCysLeuProAspP 284
|||||
1141 GTGAGTATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1190
284 rGlyGlyCysSerCysAlaThrGlyTrpLysGlyLeuGlnCysAsnGlu 300
|||||
1191 CTATGAGGATATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1240
301 AlaCysHisProGlyPheTyrGlyProAspCysLysLeuArgCysSerCy 317
|||||
1241 GATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1290
317 sAsnAsnGlyGluMetCysAspArgPheGlnGlyCysLeuCysSerProG 334
|||||
1291 TATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1340
334 tYrTrpGlnGlyLeuGlnCysGluArgGlyIleProArgMetThrPro 350
|||||
1341 GATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1390
351 LysIleValAspLeuProAspHisIleGluValAsnSerGlyLysPheAs 367
|||||
1391 CAGATAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1440
367 nProIleCysLysAlaSerGlyTrpProLeuProThrAsnGluGluMet 384
|||||
1441 CCGATATATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1490
384 hLeuValLysProAspGlyThrValLeuHisProLysAspPheCysHis 400
|||||
1491 CCGTAGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1540
401 ThrAspHisPheSerValAlaIlePheThrIleHisArgIleLeuProPr 417
|||||
1541 ACATAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1590

```















CC specification, but was obtained in electronic format directly from WIP0  
 CC at ftp.wipo.int/pub/published\_pat\_sequences.

XX Sequence 1493 BP: 425 A + 466 C + 860 G + 882 T = 0 other;

# alignment\_scores:

Quality: 31.00 Length: 453  
 Ratio: 1.891 Gaps: 14  
 Percent Similarity: 49.575 Percent Identity: 27.195

## alignment\_block:

US 09-733-764-2\_COPY\_1\_472 x AB119878/rev ..

Align seq 1/1 to reverse of: AB119878 from: 1 to: 3493

```

47 SerGlyTyrPArgProHisGluProIleThrIleThrIle..... 58
|||||..... |||
2089 TCGGATGGAAAGCAAGCGGCGAGTCCACAGACAGACAGTCCGACAGA 2040
59 ..... ArgAspPheGluAlaLeuMetAsnGlnHisG 69
||||| .....
2039 AGGATGGAAATGTGAGGAGAGAGAGAGAGAGAGAGAGAGAGAG 1990
69 InAspProLeuGluValThrGluAspValThr.....ArgGlu 81
||||| .....
1989 TGGATGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1940
82 TspAlaLysLysValValTyrLysArgGluLysAlaSerLysIleAsnG 98
||||| .....
1939 TCGATGCCAACA..... 1928
98 yAlaTyrPheCysGluGlyArgVal...ArgGlyGluAlaIleArgIleA 114
||||| .....
1927 GAGATGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1891
114 rThrMetLysMetArgGluGluAlaSerPheLeuProAlaIleThrLeu 130
||||| .....
1890 ATCAGCGGATATGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1841
131 MetThrValAspLysGlyAspAspValAsnLeuSerPheLysLysVal 147
||||| .....
1840 CGACGAGAG.....TCAGCAGCAACAGTACCGATTCGAACGCCAGCA 1794
147 IleLysGluIleAspAlaValIleTyrLys...AsnGlySerPheIle 162
||||| .....
1793 GGCATGGCGAGAGAGATGATGATGATGATGATGATGATGATGATG 1744
163 HisSerValProArgHisGluValProAspIleLeuGluValHisLeu 179
1744 ..... 1744
179 oHisAlaGlnProGluAspAlaGlyValTyrSerAlaArgTyrIleGly 196
||||| .....
1743 ...TCGAATGCCGATGACAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1708
196 LysAsnLeuPheThrSerAlaPheThrArgLeuIleValArgArgCysG 212
||||| .....
1707 .....CGATATGGAG 1696
213 ArgGlyLysTyrGlyIleValLysAsnHisLeuCysIleArgCysMet... 238
||||| .....
1695 AGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1646
229 AsnAsnGlyValGlyHisGlyAspThrGlyGluCysIleTyrPheGly 245
||||| .....
1645 GCACAGCGAGATCTGCTGACAGCGAGAGAGAGAGAGAGAGAGAGAGAG 1596
245 LysPheMetGlyArgThrCysGluLysAlaCysGluLeuHisThrPheCly 261
||||| .....
1595 GATGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1546
262 ArgThrCysLysGluArgCysSerClyGlnGluClyCysSerTyrVal 278

```

```

1545 CCGAATGGGAGTGAACCTGCAAC..... |||||..... 33
278 lPheCysLeuPheAspPheTyrGlyCysSerCysAlaIleThrIleTyrGly 295
||||| ..... ||||| ..... ||||| .....
1507 CAATGCTGCGGATGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1458
295 lPheCysLeuPheAspPheTyrGlyCysSerCysAlaIleThrIleTyrGly 311
||||| ..... ||||| ..... ||||| .....
1457 GATTAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1408
312 LysLeuArgCysSerCysAsnAsnGlyGluMetCysAspArgIleCysIle 128
||||| ..... ||||| ..... ||||| .....
1407 GCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1358
328 LysCysSerProGlyTyrGlnGlyLeuGlyCysGlyArg..... 342
||||| ..... ||||| ..... ||||| .....
1357 CCAGTGGCTTGCACGCCCGGATGCAAAACATCAAGTGGACGCCGCGG 1308
343 .....GluGlyIleProArgMetThrProLysIleValAspLeuPro 356
||||| ..... ||||| ..... ||||| .....
1307 GTGATCGAATCAATTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1258
357 AspHisIleGluValAsnSerGlyLysPheAsnProIleCysLysAlaSer 373
||||| ..... ||||| ..... ||||| .....
1257 AATAAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1211
373 rGlyTyr 375
||||| .....
1210 CGGCTGG 1204

```

seq\_name: /SIS1/gendata/geneseq/geneseq\_emb1/NA2001H.DAT:AB104034

seq\_documentation\_block:

ID: AB104034 standard; cDNA; 17478 BP.

```

XX AC AB104034;
XX DT 26-MAR-2002 (first entry)
XX DE Drosophila melanogaster expressed polynucleotide SEQ ID NO 6584.
XX KW Drosophila; developmental biology; cell signalling; insecticide;
XX KW pharmaceutical; gene; ss.
XX OS Drosophila melanogaster.
XX PN WO200171042-A2.
XX PD 27-SEP-2001.
XX PF 23-MAR-2001; 2001WO-0809231.
XX PR 23-MAR-2000; 2000US-191637P.
XX PP 11-091-2000; 2000US-0614150.
XX PA (PEKE ) PE CORP NY.
XX PI Venter JC, Adams M, Li PWD, Myers EW;
XX WK 2001-05080075.
XX DR P-PSDB; ABH59931.

```

PT New isolated nucleic acid detection reagent for detecting 1000 or more genes from *Drosophila* and for elucidating cell signalling and cell-cell interactions.

PS Claim 1; SEQ ID NO 6584; 21pp + Sequence listing; English.

CC The invention relates to an isolated nucleic acid detection reagent capable of detecting 1000 or more genes from *Drosophila*. The invention is useful in elucidating cell signalling and cell-cell interactions in higher eukaryotes for the development of





search time (sec): 51.100000

[illegible]

34 euValSerAspAlaGluThrSerLeuThrCysIleAlaSerGlyTrpArg 50  
|||||  
249 TTCTATCTGCAAGCTGGAACATCTCTCACTGCAATCCCTCTGGGCGGC 298  
51 ProHisGluProGluThrIleGlyArgAspPheGluAlaLeuMetAsnG1 67  
|||||  
299 GCGCATGAGGCAATGATATAGAGAGAGAGAGAGAGAGAGAGAGAGAG 348  
67 nHisGluAspProGluValThrGlnAspValThrArgGluTrpAlaIle 84  
|||||  
349 GCACGAGGAGGCGCGGCAAGTACTCAAGATGTCGACGAGAGAAATGGG 398  
84 yValValValTrpLysArgGlnLysAlaSerLysIleAsnGlyAlaTrp 100  
|||||  
399 AAAAAGTTGTTTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 448  
101 PheCysGluGlyArgValArgGlyGluAlaIleArgIleArgThrMetL 117  
|||||  
449 TTCTGTGAGAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 498  
117 sMetArgGluThrAlaSerPheGluProAlaThrLeuThrMetThrVal 134  
|||||  
499 GATCGCTCAACAAGCTTCCTCCGACCACTACTTTAATTAATTAATGAA 548  
134 sPlysGlyAspAsnValAsnIleSerPheLysLysValLeuIleLysGlu 150  
|||||  
549 ACAAGGAGATCAAGCTGCAATATATCTTCAAAAAGCTATGATTAAAGA 598  
151 GluAspAlaValIleThrLysAsnGlySerPheIleHisSerValProAr 167  
|||||  
599 GAAGATGCTACTATTATCAAAATGCTTCTTATCTATCTATCTATCT 648  
167 qHisGluValProAspIleGluGluValHisIleuProHisAlaGlnPro 184  
|||||  
649 GCATGAAGTACCTGATATCTGAGAGAGAGAGAGAGAGAGAGAGAGAG 698  
184 InAspAlaGlyValTrpSerAlaLeuThrIleGlyLysLeuPheThr 200  
|||||  
699 AGATGCTGAGTACTGCGGAGAGATATAGAGAGAGAGAGAGAGAGAG 748  
201 SerAlaPheThrArgLeuIleValArgAspCysGlnAlaGlySerPro 217  
|||||  
749 TCGGCTTCAAGAGCTATATATATATATATATATATATATATATATAT 798  
217 yProGluCysAsnHisIleuGlySerAlaCysMetAsnAsnGlyValCys 234  
|||||  
799 ACCTGAATGCAACCATCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAG 848  
234 IsGluAspThrGlyGluGlyLysIleCysProProGlyPheMetGlyArgThr 250  
|||||  
849 ATCAAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 898  
251 CysGlnLysAlaCysGlnAlaHisThrPheClyArgThrCysLysGluArg 267  
|||||  
899 TGTGAT 948  
267 qCysSerGlyGluGluCysLysSerLysValPheCysIleLeuProAsp 284  
|||||  
949 GTGCTGAG 998  
284 rGlyThrLysSerPysAlaThrGlyTrpLysIleLeuGlnCysAsnGlu 300  
|||||  
999 CCTATGGGCTCTCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1048  
301 AlaCysHisProGlyPheThrGlyProAspCysLysLysArgCysSerGly 317  
|||||  
1049 GCAATGAT 1098  
317 sAsnAsnGlyGluMetCysAspArgPheGlnIleCysIleuCysSerProG 334  
|||||  
1099 CAACATGGAG 1148  
334 IyTrpGlnGlyLeuIleCysGluArgGlnGlyIleProArgMetThrPro 350

1149 GATGGAGAGAGTCTAGTGTATATATATATATATATATATATATATAT 1198  
351 LysIleValAspLeuProAspHisIleGluValAsnSerGlyLysPheAs 367  
|||||  
1199 AGATATGCTATTTGATATATATATATATATATATATATATATATATAT 1248  
367 nProIleCysLysAlaSerGlyTrpProLeuProThrAsnGluGluMet 384  
|||||  
1249 TCCCATTTGCAAAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1298  
384 hrLeuValLysProAspGlyThrValLeuHisProLysAspPheAsnHis 400  
|||||  
1299 GATGCTAT 1348  
401 ThrAspHisPheSerValAlaIlePheThrIleHisArgIleLeuPro 417  
|||||  
1349 AGGATATATTTTATATATATATATATATATATATATATATATATAT 1398  
417 oAspSerGlyValTrpValCysSerValAsnThrValAlaGlyMetValG 434  
|||||  
1399 TCACTAT 1448  
434 LuLysProPheAsnIleSerValLysValLeuProLysProLeuAsnAla 450  
|||||  
1449 AAAAGGCTTCAACATTTCTGTAAAGTCTTCAAAAGGCTTCAATGCC 1498  
451 ProAsnValIleAspThrGlyHisAsnPheAlaValIleAsnIleSerSe 467  
|||||  
1499 TCAAAAT 1548  
467 rGluProTyrPheGly 472  
|||||  
1549 TGAGGCTTACTTTGGG 1564

seq\_name: /cqn\_2\_5/ptodata/1/usa/PCUS\_Comb.seq:PC1-US93-06093-1

seq\_documentation\_block:  
; Sequence 1, Application PC/TUS9306093  
; GENERAL INFORMATION:  
; APPLICANT: Giedler, Steven F.  
; TITLE OF INVENTION: NOVEL TYROSINE KINASE  
; NUMBER OF SEQUENCES: 3  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Immunex Corporation  
; STREET: 51 University Street  
; CITY: Seattle  
; STATE: Washington  
; COUNTRY: US  
; ZIP: 98101  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent's Release #1.0, Version #1.05  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PC1/US93/06093  
; FILING DATE: 19930625  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; AFFILIATION NUMBER: US 07/905,600  
; FILING DATE: 26-JUN-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Seese, Kathryn A.  
; REGISTRATION NUMBER: 32,172  
; REFERENCE/DOCKET NUMBER: 2609  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (206) 587-0430  
; TELEFAX: (206) 233-0644  
; TELEX: 756822  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 4138 base pairs



```

: TYPE: NUCLEIC ACID
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULAR TYPE: cDNA to mRNA
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 149..3523
: PCT-US93-06093-1

```

```

alignment_scores:
  Quality: 2695.00      Length: 472
  Ratio: 5.619          Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 100.000

```

```
alignment_block:
```

```
US-09-733-764-2_COPY_1_472 x PCT-US93-06093-1
```

```
Align seq 1/1 to: PCT-US93-06093-1 from: 1 to: 4138
```

```

1 MetAspSerLeuAlaSerLeuValLeuCysGlyValSerLeuLeuLeuSe 17
|||||
149 AUGGACCTCTTAGCGACCTTAGTCTCTGCGAGACAGCTGCTCTCTTC 198
|||||
17 rGlyThrValGluGluAlaMetAspLeuLeuLeuLeuLeuProL 34
|||||
194 TGGAACTGTGGCAAGGTGCGATGACCTTGATCTTGATCAATCCCTACCTC 248
|||||
34 cuValSerAspAlaGluThrSerLeuThrCysIleAlaSerGlyTrpArg 50
|||||
249 TTGATCTGATGCTTAAATATTTTAACTTAACTTAACTTAACTTAACTT 298
|||||
51 proHisGluProIleThrIleGlyArgAspPheGluAlaLeuMetAsnG 67
|||||
299 CCCCATGAGCCCATCACCATTAGTAAGGACCTTGAAGCCCTTAATCAACCA 348
|||||
67 nHisGluAspProLeuGluValThrGlnAspValThrArgGluTrpAlaL 84
|||||
349 GCACCTAGATCTCCCTGCAAGTACTCAAGATGAGACAGCAATCGGCTA 398
|||||
84 yHisValValThrPheArgGluAlaSerIleAsnGlyAlaTrp 100
|||||
399 AAAAAGTTGTTTGAACAGAGAAGGTTATGAAGATCAATGGTCTTAT 448
|||||
101 PheCysGluGlyArgValArgGlyGluAlaIleArgIleArgThrMetI 117
|||||
449 TTCCTGAAAGGGCGAGTTCGAGGAGAGGCAATCAGATACGAACCATGAA 498
|||||
117 sMetArgGlnGlnAlaSerPheLeuProAlaThrLeuThrMetThrVal 134
|||||
499 GATCGGTCACCAAGCTTCCTTCCCTACCAGCTACTTTAACTATGACTGTG 548
|||||
134 sPlysGlyAspAsnValAsnIleSerPheIleValIleValIleValI 150
|||||
549 ACAAGGAGATATAATGTAATATATCTTTTAAAAGGATATGATTAAGAA 598
|||||
151 GluAspAlaValIleThrValAsnGlySerPheIleHisSerValProAr 167
|||||
599 GAACATCCATGATTTACAAAAATGGTTCCTCATCCATTCAGTCCGCCG 648
|||||
167 qHisGluValProAspIleLeuGluValHisLeuProHisAlaGlnProG 184
|||||
649 GCAATGAAGTACCTGCAATATATTAAGATACCTGCGGCACTGCACTGCT 698
|||||
184 InAspAlaGlyValThrSerAlaArgThrIleCysGlyAsnIleCysThr 200
|||||
699 AGGATGTGAGTGTATATGAGTATATATATATATATATATATATATATAT 748
|||||
201 SerAlaPheThrArgGluLeuValArgGlyCysGluAlaGluIleTrpG 217
|||||
749 TCGGCTTCACCAAGGTCGATGATGATGATGATGATGATGATGATGATGAT 798
|||||
217 yProGluCysAsnHisLeuCysThrAlaCysMetAsnGlyValCysH 234

```

```

|||||
799 ACCTGAATGCAACCATCTCTGACTGCTGTATGAACAATGGTCTGGC 848
|||||
234 TSCThAspThrGlyCysIleCysProProGlyPheMetClyArgThr 259
|||||
849 ATGAAGATATGAGAAATGATTTTGGCTGCTGAGGATATGAGGAAAG 898
|||||
251 CysGluLysAlaCysGluLeuHisThrPheGlyArgThrCysLysGluAr 267
|||||
899 TGTGAGAAAGGCTTGTCACTGACACGCTTGGCAGAACCTGTAAAGAAAG 948
|||||
267 qCysSerGlyGluGluCysLysSerTyrValPheCysLeuProAspF 284
|||||
949 GTTATATGACAAAGGATATGAGTCTTATGAGTCTGAGTCTGAGTCT 998
|||||
284 rGlyThrCysSerCysAlaThrGlyThrPheCysGluCysAsnGlu 309
|||||
999 CTTATGAGGCTTGTGAGTATGAGTATGAGTATGAGTATGAGTATGAG 1048
|||||
301 AlaCysHisProGlyPheThrGlyProAspCysLysLeuArgCysSerCy 317
|||||
1049 GCATGCCACGCTGCTTAAAGGCTGACATGCTAAAGCTTAAAGCTGAC 1098
|||||
317 sAsnAsnGlyGluMetCysAspArgPheGlnGlyCysLeuCysSerProG 334
|||||
1099 CAAATATGAGGCTATGAGTATGAGTATGAGTATGAGTATGAGTATGAG 1148
|||||
334 LyrPheGlnGlyLeuGlnCysGluArgGlyIleProArgMetThrPro 350
|||||
1149 GATGATGAGGCTTGTGAGTATGAGTATGAGTATGAGTATGAGTATGAG 1198
|||||
351 LysIleValAspLeuProAspHisIleGluValAsnSerGlyLysPheAs 367
|||||
1199 AAGATAGTGCATTTGCCACATCATATAGCAAGTAAACAGTGGTAAATTA 1248
|||||
367 nProIleCysLysAlaSerGlyTrpProLeuProThrAsnGluGluMet 384
|||||
1249 TCCATATATGAGGCTTGTGAGTATGAGTATGAGTATGAGTATGAGTAT 1298
|||||
384 hrLeuValLysProAspGlyThrValLeuHisProLysAspPheAsnHis 400
|||||
1299 CCTGTGTGAAGTGTGAGTATGAGTATGAGTATGAGTATGAGTATGAGT 1348
|||||
401 ThrAspHisPheSerValAlaIlePheThrIleHisArgIleLeuProPr 417
|||||
1349 ACCGATCATTTTCAGTAGGCTATATTCACCTATCCACCTGCAAGGAGT 1398
|||||
417 oAspSerGlyValTrpValCysSerValAsnThrValAlaGlyMetValG 434
|||||
1399 TGAATATGAGGCTTGTGAGTATGAGTATGAGTATGAGTATGAGTATGAG 1448
|||||
434 lLysProPheAsnIleSerValLysValIleProLysProLeuAsnAla 450
|||||
1449 AAAAGGCTTGAATATTTTGTAAATGCTTGAAGGCTTGAAGGCTTGAAG 1498
|||||
451 ProAsnValIleAspThrGlyHisAsnPheAlaValIleAsnIleSerSe 467
|||||
1499 CCAACCTGATTCACACTGCACATACTTTGCTGCTATCAACATCAGCTC 1548
|||||
467 rGluProTyrPheGly 472
|||||
1549 TGAAGCTTATTTTGGG 1564

```

```
seq_name: /qgn2_6/ptodata/1/lna/5A_00M8 seq-US-07-934-393B-1
```

```
seq_documentation_block:
```

```
: Sequence 1. Application US/07934393B
```

```
: Patent No. 5466596
```

```
: GENERAL INFORMATION:
```

```
: APPLICANT: BREITMAN, MARTIN L.
```

```
: APPLICANT: DOMONT, DANIEL
```

```
: APPLICANT: GRADWIL, GERAUD G.
```

```
: TITLE OF INVENTION: TISSUE SPECIFIC TRANSCRIPTIONAL
```





167	qhlisgluValProAspIleLeuGluValHisLeuProHisAlaGlnProG	184
624	GCAACAACTATCTGATATTTTACAAGTTTCACTACCCGCTACGG	673
184	InAspAlaGlyValTyrSerAlaArgTyrIleGlyGlyAsnLeuPheThr	200
674	AGGATGCTGTGTTGTTATTCGGGTAAGTATATAGAGAGAAATGTT	723
201	SerAlaPheCysArgGluLeuValTargArgCysGluAlaGlnTyrPhe	217
724	TCCAGCTTCACCAAGCTGATTTTCGAGAGATGTAAGATCAAGATG	773
217	ProGluCysAsnHisLeuCysThrAlaCysMetAsnAsnGlyValCys	234
774	GGCGACATGACGCTGCTGCTATGATACATTCACCAATCAATGGAC	823
234	AspGluAspThrCysGlyCysIleCysProGlnTyrPheMetGlyArg	250
824	ATGAGATATATGAGAAATGATTTGATGTTTGGTTTATGGGAAATA	873
251	CysCysIleValCysCysGluCysHisThrPheCysArgThrCysIle	267
874	TGTTAGAAAGATTCTGAGAGGCAATATTTTCTCAAGATTTGAA	923
267	AspSerGlnGlnGlnGlyCysLysSerTyrValPheCysLeuProAsp	284
924	GATGACGACGACCAAGACATGCAAGCTATATGCTGTCGCTGCTCC	973
284	ProTyrGlyCysSerGlyAsnTyrGlyTyrLysGlyLeuGlnGly	300
974	CTTAAGGCTGTTCTGTGCTAAAGATGAGAGGATTTGTAAGAAAG	1023
301	AlaCysHisProGlyPheTyrGlyTyrAspCysLysLeuAspCysSer	317
1024	GGATGCTATATGATTATATAGAGATATATATGTAAGATATATGTT	1073
317	AsnAsnGlyValMetCysAspArgPheGlnGlyCysLeuCysSerPro	334
1074	TACCAATCAACAGATATGTCATTCCTCTTCAAGCAATGCTGCTCA	1123
334	TyrTrpGlnGlyLeuGlnCysGlnArgGlnGlyIleProArgMetThr	350
1124	GATGCAAGAGGCTGCTATGTCTCAAAATGAGAGAGCAATGATGCA	1173
351	LysIleValAspIleuProAspHisIleGluValAsnSerGlyLysPhe	367
1174	CAGATAGAGAGATTTGGAGATCAATATGAGATAAAGTGGAAATTTAA	1223
367	ProIleCysLysAlaSerGlyTyrPheLeuProThrAsnGlnGluMet	384
1224	CGCAATCTCAACAGCCGCTGGGCGCACTACCTACAGTCAAGAAATGA	1273
384	hrLeuValLysProAspGlyThrValLeuHisProLysAspPheAsnHis	400
1274	GGCTAGTCAAGAGATATGAGAAATGCTGCAAGATGAGATTTCAATAT	1323
401	ThrAspHisPheSerValAlaIlePheThrIleHisArgIleLeuPro	417
1324	ACAGATCGTTCTTCAGTGGCAATATTCAGTGCACAAAGATCTTACCT	1373
417	AsnSerGlyValTrpValCysSerValAsnThrValAlaGlyMetValG	434
1374	TGACTCAAGCAGTCGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1423
434	IuLysProPheAsnIleSerValLysValLeuProLysProLeuAsnAla	450
1424	AAAGAGCTTTCACATTTCTGCTCAAAATTTCTTCTCAGAGCTCTGACAG	1473
451	ProAsnValIleAspThrGlyHisAsnPheAlaValIleAsnIleSerSe	467
1474	GCAATATGATTTGATATGCAATATATTTTGTATATATATATATATAT	1523

Align seq 1/1 to: US-08-938-957A-1 from: 1 to: 4175

```

1 MetAspSerLeuAlaSerLeuValLeuCysGlyValSerLeuLeuLeu 17
124 ATGGACTCTTTAGACGGTCTAGTCTCTGTGTGATCTTAGCTTGTCTTAA 173
17 rGlyThrValGluGlyAlaMetAspLeuLeuLeuLeuLeuLeuLeu 34
174 TGGAGTATAGAAAGGACATGAGGAGATGATCTTATGATCTTATGATCT 223
34 euValSerAspAlaGluThrSerLeuThrCysLeuAlaSerGlyTyrArg 50
224 TTGTGTCTCATGCCGAACATCCCTCCACCTCCATGCCCTCGGTGCCAC 273
51 ProHisGluProIleThrIleGlyArgAspPheGluAlaLeuMetAsnG 67
274 GCGATATATATATATATATATATATATATATATATATATATATATAT 323
67 rHisGlnAspProLeuGluValThrGlnAspValThrArgGluTrpAla 84
324 GCACCAAGATCCTGAGAGTATCTAAGATGAGTACAGAGATGAGAGAG 373
84 yLysValValTrpLysArgGluLysAlaSerLysIleAsnGlyAlaTyr 100
374 AAAAAAGTCTTTGGAAGAGAGAAAGGCCAGTAAGATTAATGCTGTAT 423
101 PheCysGluGlyArgValArgGlyGluAlaIleArgIleArgThrMetLy 117
424 TTCTGTGAAGTCTGAGTTCGAGGAGATAGGCTATATAGGATACGAG 473
117 sMetArgGlnGlnAlaSerPheLeuProAlaIleLeuThrMetThrVal 134
474 CATGATCTTAAATATATATATATATATATATATATATATATATATAT 523
134 sPlySGlyAspAsnValAsnIleSerPheLysLysValLeuIleLysGlu 150
524 ACAGGGAGATATCTGAACATATCTTTCAAAAGGCTGTATATTAAGAA 573
151 GluAspAlaValIleTyrLysAsnGlySerPheIleHisSerValProAr 167
574 GAAGATGAGGATATATCAAAATATGGCTCTCTATCCACTACGTCGCC 623
167 rHisGluValProAspIleLeuGluValHisLeuProHisAlaGlnPro 184
624 CATGATCTTAAATATATATATATATATATATATATATATATATATAT 673
184 InAspAlaGlyValTyrSerAlaArgTyrIleGlyGlyAsnLeuPheThr 200
674 AGCATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 723
201 SerAlaPheThrArgLeuIleValArgCysGluAlaGlnIleYstPgl 217
724 TCAATCTTCAATATATATATATATATATATATATATATATATATATAT 774
217 yProGluCysAsnHisLeuCysThrAlaCysMetAsnAsnGlyValCysH 234
774 GCGGATCTGATATATATATATATATATATATATATATATATATATATAT 824
234 sGluAspThrGlyGlyGlyLysProProGlyPheMetGlyArgThr 250
824 ATGAAGATACCGGAGATATATATATATATATATATATATATATATATAT 873
251 CysGluLysAlaValGlyLeuLeuHisIleThrPheArgThrCysLysG 267
874 TCTAGAAAGCTTCTGAGAGTCAATATATATATATATATATATATATAT 924
267 rCysSerGlyGlnGluGlyCysLysSerTyrValPheCysLeuProAsp 284
924 ATGATATGAGGATATATATATATATATATATATATATATATATATATAT 974
284 rCysGlyCysSerCysAlaThrGlyTrpLysGlyLeuGluCysAsnGlu 306
974 CTTCAGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1023

```

```

301 AlaCysHisProGlyPheTyrCysProAspCysLysLeuArgCysSerCy 317
1024 GATATATATATATATATATATATATATATATATATATATATATATAT 1073
317 sAsnAsnGlyGluMetCysAspArgPheGlnGlyCysLeuCysSerProG 334
1074 TACCAATCAACAGATATCTATCTGATCTGATCTGATCTGATCTGATCT 1123
334 LyTrpGlnGlyLeuGlnCysGluArgGluGlyIleProArgMetThrPro 350
1124 GATGGCAAGGCTCTGAGTGTGATGATGATGATGATGATGATGATGAT 1173
351 LysIleValAspLeuProAspHisIleGluValAsnSerGlyLysPheAs 367
1174 CAGATAGAGGATTTGCTGATATATATATATATATATATATATATATAT 1223
367 nProIleCysLysAlaSerGlyTyrProLeuProThrAsnGluGluMet 384
1224 CCCATCTGCCAAAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1273
384 hrLeuValLysProAspGlyThrValLeuHisProLysAspPheAsnHis 400
1274 CCCIAGTCAAGCCAGATCGGACAGTCTCCACCAACCAATGACTTCACTAT 1323
401 ThrAspHisPheSerValAlaIlePheThrIleHisArgIleLeuPro 417
1324 ACAGATGCTTCTGATGATATATATATATATATATATATATATATATAT 1373
417 oAspSerGlyValTrpValCysSerValAsnThrValAlaGlyMetVal 434
1374 TCACACAGAGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1423
434 LuLysProPheAsnIleSerValLysValLeuProLysProLeuAsnAla 450
1424 AAAAGCTTTTCAACATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1473
451 ProAsnValIleAspThrGlyHisAsnPheAlaValIleAsnIleSer 467
1474 CCAAAAGTATGATGATGATGATGATGATGATGATGATGATGATGATG 1523
467 rGluProTyrPheGly 472
1524 TCAGCTTACTTTGG 1539

```

seq name: /cnp2\_5/plodata/1/ina/CA\_COMP seq: US-08-278-089A-5

seq documentation block:

: Sequence 5, Application US/08278089A

: Patent No. 5681714

: GENERAL INFORMATION:

: APPLICANT: Breitman, Martin L.

: APPLICANT: Rossant, Janet

: APPLICANT: Dumont, Daniel J.

: APPLICANT: Yamaguchi, Terry P.

: TITLE OF INVENTION: No. 5681714e1 Receptor Tyrosine Kinase

: NUMBER OF SEQUENCES: 33

: CORRESPONDENCE ADDRESS:

: ADDRESSEE: Bereskin & Parr

: STREET: 40 King Street West

: CITY: Toronto

: STATE: Ontario

: COUNTRY: Canada

: ZIP: M5H 3Y2

: COMPUTER READABLE FORM:

: MEDIUM TYPE: Floppy disk

: COMPUTER: IBM PC compatible

: OPERATING SYSTEM: PC-DOS/MS-DOS

: SOFTWARE: Patent In Release #10, Version #1.30

: CURRENT APPLICATION DATA:

: APPLICATION NUMBER: US/08/278,089A

: FILING DATE: 20-JUL-1994

: CLASSIFICATION: 530

ATTORNEY/AGENT INFORMATION:  
 NAME: Kurydyk, Linda M.  
 REGISTRATION NUMBER: 44,971  
 REFERENCE/DOCKET NUMBER: 3153-111  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (416) 364-7311  
 TELEFAX: (416) 361-1598  
 INFORMATION FOR SEQ ID NO: 5:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 4176 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: cDNA  
 HYPOTHETICAL: NO  
 ANTI-SENSE: NO  
 FRAGMENT TYPE: N-terminal  
 ORIGINAL SOURCE:  
 ORGANISM: Mus musculus  
 STRAIN: CD-1  
 DEVELOPMENTAL STAGE: Embryo  
 TISSUE TYPE: Heart  
 IMMEDIATE SOURCE:  
 CLONE: Tek  
 POSITION IN GENOME:  
 CHROMOSOME/SEGMENT: 4  
 MAP POSITION: Between the brown and pmv-23 loci  
 FEATURE:  
 NAME/KEY: CDS  
 LOCATION: 124...1490  
 US 08 278-089A-5

alignment\_scores:  
 Quality: 2407.00 Length: 472  
 Ratio: 5.233 Gaps: 0  
 Percent Similarity: 97.458 Percent Identity: 91.949  
 alignment\_block:  
 US-09-733-764-2\_Copy\_1\_472 x US-08-278-089A-5  
 Align seq 1/1 to: US-08-278-089A-5 from: 1 to: 4176  
 1 MetAspSerLeuAlaSerLeuValLeuCysGlyValSerLeuLeuLeuLeu 17  
 124 ATGCACTCTTTAGCGAGCTTAACTTCTCTGAGTCACTTCTCTTTA 173  
 17 rGlyThrValGluGlyAlaMetAspLeuLeuLeuLeuLeuLeuLeu 34  
 174 TGCAGTACTAGAGGCTGATGGAAGCTGATGATCAATTCCTTACGTC 223  
 34 euValSerAspAlaGluThrSerLeuThrCysIleAlaSerGlyTrpArg 50  
 224 TTGTGTCTGATGCGCAAGACATCCCTACCTGCTGCTGCTGCTGCTG 273  
 51 ProHisGluProIleThrIleGlyArgAspPheGluAlaLeuMetAsnG 67  
 274 CCGCATGAGCCATACCAATAGAGAGAGCTTGAAGGCTTAATGAAACA 323  
 67 nHisGlnAspProLeuGluValThrGlnAspValThrArgGluTrpAla 84  
 324 GCACCAAGAATCCACTGGAGGTTTACCAAGATGTGACCAAGAAATGG 373  
 84 yValValValTrpLysArgGluLysAlaSerLysIleAsnGlyAlaTrp 100  
 374 AAAAATTGTTGGAAGACAGAAAGAGCCAGTAAGATTAAATGCTGCTAT 423  
 101 PheCysGluGlyArgValArgGlyGluAlaIleArgIleArgThrMetL 117  
 424 TTCTGTGAAGTGGAGTTTGAAGGAGAGGCTATTAAGGATACGACCAT 473  
 117 sMetArgGlnGlnAlaSerPheLeuProAlaThrLeuThrMetThrVal 134  
 1374 TGCACAGAGCTCTGGTCTCTGATGTCGAAACACAGTGTCTGGGATGG 1424

474 GATCCCTCAACCAAGCATCCTCTCTACCTGCTACTTTAACTAAGACGG 524  
 134 sPlyGlyAspAsnValAsnIleSerPheLysLysValIleIleLysGlu 150  
 524 ACAGGGAATATATGGAATATATTTTCAAAAAAGCTTTAATTAAGAA 574  
 151 GluAspAlaValIleIleTrpLysAsnGlySerPheIleHisSerValPro 167  
 574 GAAGATCATGATTTACAAAAATGCTGCTCATCATCATCATCATCAT 624  
 167 qHisGluValProAspIleLeuGluValHisLeuProHisAlaGlnPro 184  
 624 GCAATGAATGCTGATATTTTACAACTTCACTTGGCGCATGCTCAGG 674  
 184 InAspAlaGlyValTrpSerAlaArgTrpIleGlyGlyAsnLeuPheThr 200  
 674 AGGATGCTGATGATGATGATGATGATGATGATGATGATGATGATG 724  
 201 SerAlaPheThrArgLeuIleValArgArgCysGluAlaGlnIleLys 217  
 724 TCAGGCTTCACCAAGGCTGATTTGTTGAGAGATGTAAGGCTCAAG 774  
 217 yProGluCysAsnHisLeuCysThrAlaCysMetAsnAsnGlyValCys 234  
 774 GCGGACTGTAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 824  
 234 IscLeAspThrGlyGluCysIleCysProProGlyPheMetGlyArgThr 250  
 824 ATCAATATATGAGGAAATGATTTGCTGCTGCTGCTGCTGCTGCTG 874  
 251 CysGluLysAlaCysGluLeuHisThrPheGlyArgThrCysLysGlu 267  
 874 TGTGAGAAATGTTGTAGGCTGATATTTGCTGCTGCTGCTGCTGCT 924  
 267 gCysSerGlyGlyGlyGlyCysLysSerTrpValPheCysLeuTrpAs 284  
 924 GGTGATGGAACCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 974  
 284 rGlyGlyCysSerCysAlaThrGlyTrpLysGlyLeuGlyCysAsnGlu 300  
 974 CTATGAGGTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1024  
 301 AlaCysHisProGlyPheTrpGlyProAspCysLysCysArgCysSer 317  
 1024 GCATGCCATCTGATATATATATATATATATATATATATATATAT 1074  
 317 sAsnAsnGlyGluMetCysAspArgPheGlnGlyCysLeuCysSerPro 334  
 1074 TACTAATCAATGATAATGATGATGATGATGATGATGATGATGAT 1124  
 334 LyrTrpGlnGlyLeuGlnCysGluArgGlnGlyIleProArgMetThr 350  
 1124 GATGGCAAGGCTGATGATGATGATGATGATGATGATGATGATGAT 1174  
 351 LysIleValIleAspLeuProAspHisIleGluValIleAsnSerGly 367  
 1174 CAAATAGAGAGATTTTAAATATATATATATATATATATATATAT 1224  
 367 nProIleCysLysAlaSerGlyTrpProLeuProThrAsnGluLeuMet 384  
 1224 CCGCATGCTGCAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1274  
 384 hrLeuValLysProAspGlyThrValLeuHisProLysAspPheAsn 400  
 1274 CTTATCTAAAGGATGATGATGATGATGATGATGATGATGATGAT 1424  
 401 ThrAspHisPheSerValAlaIlePheThrIleHisArgIleLeuPro 417  
 1324 ACAGATGCTTTCTGATGCTGATATATATATATATATATATATAT 1474  
 417 oAspSerGlyValTrpValCysSerValAsnThrValAlaGlyMetVal 444  
 1374 TGCACAGAGCTCTGGTCTCTGATGTCGAAACACAGTGTCTGGGAT 1424



```

874 TGTGAGAAAGTTGTGAGAGGAGTAAATTTGGAGAGTGTAAAGAAAG 923
267 gCysSerGlyGlnGluCysSerSerValThrCysLeuProAsp 284
|||||
924 GTGTAGTGCACAGAGCAAGCAAGCTTATGTGTCTGTCTCCAGCC 973
284 rGlyGlyCysSerCysAlaThrGlyTrpLeuGlyLeuGlyCysAsnGlu 300
|||||
974 GTTACGGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1023
401 AlaCysHisThrGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 317
|||||
1024 GATGCGCAATCTGGTATATAGAGAGAGCTTAAAGTCAAGTCAATG 1073
317 sAsnAsnGlyCysMetCysAspThrGlyGlyGlyGlyGlyGlyGlyGly 334
|||||
1074 TACCAATGAAGACATAATGATGCTTCAAGGATGCTCTCTCTCTCTCT 1123
334 lyTrpGlnGlyLeuGlnCysGluArgGluGlyIleProArgMetThrPro 350
|||||
1124 GATGGCAAGCGCTGAGCTCTCAGAAAGCAAGGACGCGCAAGGATGAC 1173
351 lysIleValAspLeuProAspHisIleGluValAsnSerGlyLysPheAs 367
|||||
1174 CAGATGAGCATTTTGTATATATATATATATATATATATATATATAT 1223
367 nProIleCysLysAlaSerGlyTrpProLeuProThrAsnGluMet 384
|||||
1224 CCGCATCTGCAAGCGCTGAGCTCTCAGAAAGCAAGGACGCGCAAGGAT 1273
384 hrLeuValLysProAspGlyThrValLeuLysProLysAspPheAsnHis 400
|||||
1274 CCTAGTCAAGCCAGATGGCAGACAGTGGCCCAAGCAAAATCACTTCA 1323
401 ThrAspHisPheSerValAlaIlePheThrIleHisArgIleLeuPro 417
|||||
1324 ATAGATGTTTCTATATATATATATATATATATATATATATATATAT 1373
417 oAspSerGlyValTrpValCysSerValAsnThrValAlaGlyMetVal 434
|||||
1374 TGACTCAGCAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1423
434 hLysProPheAsnIleSerValLysValLeuProLysProLysAsnAla 450
|||||
1424 AAAGCGTTTCAACATTCGGCGCAAGTCTTCCAGAGCGCGCTGCAGCC 1473
451 proAsnValIleAspThrGlyHisAsnProAlaValIleAsnIleSer 467
|||||
1474 CCAATGTCATTCATATATATATATATATATATATATATATATATAT 1523
467 rGluProLysPheGly 472
|||||
1524 TGAGCGCTACCTTTGGG 1549

```

seq name: /cseq2\_6/ptdata/17/seq/56\_0\_M0.seq.us 08 220-240A-4

seq documentation\_block:

```

: Sequence 4, Application US/08220240A
: Patent No. 5955291
: GENERAL INFORMATION:
: APPLICANT: Allitalo, Karl
: APPLICANT: Matikainen, Marja-Terttu
: APPLICANT: Partanen, Juhani
: APPLICANT: Makela, Tomi
: APPLICANT: Korhonen, Jaana
: TITLE OF INVENTION: ANTIBODIES RECOGNIZING TIE RECEPTOR
: TITLE OF INVENTION: TYROSINE KINASE AND USES THEREOF
: NUMBER OF SEQUENCES: 5
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Marshall, U'loole, Gerstein, Murray & Borun
: STREET: 231 South Wacker Drive, 4000 South Tower
: CITY: Chicago
: STATE: Illinois

```

```

: COUNTRY: United States of America
: ZIP: 60606-6402
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.40
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/220,240A
: FILING DATE: 29-MAR-1994
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: PC/FT94/00006
: FILING DATE: 08-JAN-1993
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/817,800
: FILING DATE: 09-JAN-1992
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/167,453
: FILING DATE: 15-DEC-1993
: ATTORNEY/AGENT INFORMATION:
: NAME: Gass, David A.
: REGISTRATION NUMBER: 38,153
: REFERENCE/DOCKET NUMBER: 29151/31958
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (312) 474-6300
: TELEFAX: (312) 474-0448
: INFORMATION FOR SEQ ID NO: 4:
: SOURCE CHARACTERISTICS:
: LENGTH: 3845 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 37..3450
: US-08-220-240A-4

```

alignment\_scores:

Quality:	865.50	Length:	454
Ratio:	2.847	Gaps:	7
Percent Similarity:	66.960	Percent Identity:	37.885

alignment\_block:

US-09-733-764-2\_copy\_1\_472 x US-08-220-240A-4

Align: seq 1/1 to: US-08-220-240A-4 from: 1 to: 4845

```

14 LeuLeuLeuSerGlyThrValGluGlyAlaMetAspIleLeuLeuAs 30
|||||
73 GTGTCTGTGGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 122
48 nSerLeuProLeuValSerAspAlaGluThrSerLeuThrCysIle 45
|||||
123 CAACCTGGCGCTCAGGACGCCGCCCTTCTCTCTCTCTCTCTCTCTCTCT 172
46 .....AlaSerGlyTrpArgProHisGluPro 54
|||||
173 GATATATATATATATATATATATATATATATATATATATATATATAT 216
55 IleThrIleGlyArgAspPheGluAlaLeuMetAsnGlnHisGlnAsp 71
|||||
217 GTGGCTGGCTGGACCAACCAACCAACCAACCAACCAACCAACCAACCA 266
71 LeuGluValThrGlnAspValThrArgGluTrpAlaLysLysValVal 88
|||||
247 CTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 315
88 TrpLysArgGlyLysAlaSerLysIleAsnGlyAlaTrpPheGlySer 104
|||||
316 .....TCTAAAGCTCTGAGCTCTCTCTCTCTCTCTCTCTCTCTCT 357

```







```

alignment_scores:
  Quality: 231.50      Length: 248
  Ratio: 2.086        Gaps: 14
  Percent Similarity: 44.758  Percent Identity: 26.613

alignment_block:
  US-09-733-764-2_COPY_1_472 x US-08-872-855-3 from: 1 to: 2055
  Align seq 1/1 to: US-08-872-855-3 from: 1 to: 2055

179 ProHisAlaGlnProGlnAspAla      GlyValTyrSe 190
|||||
391 CCAGAGGCTTGGCCAGAGATGGCACTCATCAGCAAGATGCCATCCAGGG 440
|||||
190 rAlaArgTyrIleGlyGlyAsnLeuPhe      ThrSerAlap 203
|||||
441 CTCCTAGCTGTGGTACAGAACTGTTATTGATATGAAATGAAGTAAAGTAA 400
|||||
203 beThrArgLeu.....IleValArgArgCysGluAlaGln 214
|||||
491 TCACAAGGCTCGCTACTCTTACCGGAGGAGATGAGTGGCAGTGGACAA 534
|||||
215 LysArgPheProGlyCysAsnHisLeuPysThr      255
|||||
535 TACTATGAGACAACTCTCCGCGCTGTGTGTAAGTAAGGCGCAATGACCACTT 584
|||||
226 ..... AlaCysMetAsnAsnGly      231
|||||
585 GGGTCACTATGTGTGTAAGTAAGTAAATGTAAGTAAATGTTGCTTATGAGTT 634
|||||
232 ..... ValCysHisGln 235
|||||
635 GGACTGGGGAATATTGCCAAGAGCCTATCTCTTCGGGCTGTGCATGAA 684
|||||
236 AspThrGly      GluCysIleCysProGlyGlyPhe 246
|||||
685 CAGAAATGGCTACTGCCAGCAAGCCAGCAGAGTGTCTCTGCCGCCAGGCGTG 734
|||||
246 cMetClyArgThrCysGluIysAlaCysGluLeuHisThrPheGlyArgT 263
|||||
735 GCAGGGCGGGTGTGTAAAGAA      756
|||||
263 hrCysLysGluArgCysSerGlyGlnGluClyCysLysSerTyrValPhe 279
|||||
757 TGCATGCGCCACAAATGGCTGTCTGCC...CAGGGCAACC 789
|||||
280 CysLeuProAspProTyrGlyCysSerCysAlaThrGlyTyrLysGlyLe 296
|||||
790 TGC AGCACTGCTGCGCAAAAGTACTGTGTGATGAGGGGCTGGGGAGGCGCT 836
|||||
296 hClnCysAsnGln      300
|||||
837 GTTTGTGAGCAAAATATTAATATATGAAATATATATATATATATATATATAT 886
|||||
301 ..... AlaCys 302
|||||
887 ATGGGGCAACCTGCTTCAACAGTGGTATGCAATCAATCAATCAATCAATCAAT 936
|||||
303 HisProGlyPheTyrGlyProAspHisCysLysLeuArgCysSer 316
|||||
937 GCGCCAGAGCTACACTGGTGTGTGATGCAATCAATCAATCAATCAATCAATCA 986
|||||
317 ..... CysAsnAsnGlyGlnMetCys      AspArgPheClnGly 329
|||||
987 CAGTAAGGCTATGTGTAAGTAAGTAAATGTAAGTAAATGTAAGTAAATGTAAG 1036
|||||
329 ..... CysLeuCysSerSerSerGlyThrGlnGlyLeuGlnCysGln 341
|||||
1037 ACCACTGCTGTGTGCTGGGGGCTACTATATATATATATATATATATATATAT 1080
|||||

seq_name: /can2_6/prodat3/1/109/8A/00MB seq-us-08-872-855-1

```

```

seq_documentation_block:
  Sequence 1, Application US/08872855
  Patent No. 6:21045
  GENERAL INFORMATION:
  APPLICANT: McCarthy, Sean
  APPLICANT: Gearing, David
  TITLE OF INVENTION: NOVEL HUMAN DELTA3 COMPOSITIONS AND
  TITLE OF INVENTION: THERAPEUTIC USES THEREFOR
  NUMBER OF SEQUENCES: 23
  CORRESPONDENCE ADDRESS:
  ADDRESSEE: FOLEY, HUAG & ELLIOT LLP
  STREET: One East Office Square
  CITY: Boston
  STATE: MA
  COUNTRY: USA
  ZIP: 02109-2170
  COMPUTER READABLE FORM:
  MEDIUM TYPE: Floppy disk
  COMPUTER: IBM PC compatible
  OPERATING SYSTEM: PC-DOS/MS-DOS
  SOFTWARE: Patent In Release #1.0, Version #1.30
  CURRENT APPLICATION DATA:
  APPLICATION NUMBER: US/08/872,855
  FILING DATE: 11-JUN-1997
  CLASSIFICATION: 514
  ATTORNEY/AGENT INFORMATION:
  NAME: Arnold, Beth E.
  REGISTRATION NUMBER: 35,430
  REFERENCE/DOCKET NUMBER: MAA-603.02
  TELECOMMUNICATION INFORMATION:
  TELEPHONE: 617-832-1000
  TELEFAX: 617-832-7590
  INFORMATION FOR SEQ ID NO: 1:
  SEQUENCE CHARACTERISTICS:
  LENGTH: 2800 base pairs
  TYPE: nucleic acid
  STRANDEDNESS: single
  TOPOLOGY: linear
  MOL. CULE TYPE: cDNA
  FEATURE:
  NAME/KEY: CDS
  LOCATION: 338..2392
  US-08-872-855-1

alignment_scores:
  Quality: 231.50      Length: 248
  Ratio: 2.086        Gaps: 14
  Percent Similarity: 44.758  Percent Identity: 26.613

alignment_block:
  US-09-733-764-2_COPY_1_472 x US-08-872-855-1 from: 1 to: 2860
  Align seq 1/1 to: US-08-872-855-1 from: 1 to: 2860

179 ProHisAlaGlnProGlnAspAla      GlyValTyrSe 190
|||||
178 CCAGAGGCTTGGCCAGAGATGGCACTCATCAGCAAGATGCCATCCAGGG 777
|||||
190 rAlaArgTyrIleGlyGlyAsnLeuPhe      ThrSerAlap 203
|||||
176 CTCCTAGCTGTGGTACAGAACTGTTATTGATATGAAATGAAGTAAAGTAA 827
|||||
203 beThrArgLeu.....IleValArgArgCysGluAlaGln 214
|||||
828 TCACAAGGCTCGCTACTCTTACCGGAGGAGATGAGTGGCAGTGGACAA 871
|||||
215 LysArgPheProGlyCysAsnHisLeuPysThr      255
|||||
872 TACTATGAGACAACTCTCCGCGCTGTGTGTAAGTAAGGCGCAATGACCACT 921
|||||
226 ..... AlaCysMetAsnAsnGly      231

```





```

alignment_scores:
  Quality: 209.50      Length: 322
  Ratio: 1.406         Gaps: 20
  Percent Similarity: 46.273      Percent Identity: 24.345

alignment_block:
US-09_733-764-2_COPY_1_472 x US-08-949-366-27 ..
  Align seq 1/1 to: US 08-949-366-27 from: 1 to: 6677

211  CYSCLALALALALysrpsdlyrprochcysasrhls..... 222
2172  TGTGATGACATGTCATCTGACATTCACATGACATGACACACAACTGGCTGTGGCTC 2421
223  LeuGysTrpAlaGlyGlyMetAsnAsnGlyValGlySerHis..GluAspSerG 238

```

[illegible]

3268 CAACCAAGGAA 3279

---









```

1 MetAspSerLeuAlaSerLeuValLeuCysGlyValSerLeuLeuLeu 17
|||||
86 ATGAGATTTTATGCGGTTAA:TT-T-T-T-T-T-T-T-T-T-T-T-T-T-T 135
|||||
17 rClyThrValGluGlyAlaMetAspLeuLeuLeuLeuLeuLeuLeu 34
|||||
136 TGGAGTACAGAGCGGCGGATGAGCGGATTTTATGATCAATTCCTACCTC 185
|||||
34 euValSerAspAlaGluThrSerLeuThrCysIleAlaSerGlyTrpArg 50
|||||
186 TGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 245
|||||
51 ProHisGluProIleThrIleGlyArgAspPheGluAlaLeuMetAsnG 67
|||||
236 CCGCATATGCGATATATATATATATATATATATATATATATATATAT 295
|||||
67 nHisGlnAspProLeuGluValThrGlnAspValThrArgGluTrpAla 84
|||||
286 ATACCAAGATCATGAGAGGTTACTTCAAGATGAGGAGAGATGCGGCA 335
|||||
84 yLysValValTrpLysArgGluLysAlaSerLysIleAsnGlyAlaTyr 100
|||||
336 AAAAACTTCTTGGGAAGAGAGAAAGGCGGATTAAGATTAAGTGGCTAT 385
|||||
101 pheCysGluGlyArgValArgGlyGluAlaIleCargIleArgThrMet 117
|||||
386 TT-T-T-T-T-T-T-T-T-T-T-T-T-T-T-T-T-T-T-T-T-T-T-T-T-T 445
|||||
117 sMetArgGlnGlnAlaSerPheLeuProAlaThrLeuThrMetThrVal 134
|||||
436 GATGCGCTCAACAGCATCTCTTACCGCTGCTTAACTATGACCGTGG 485
|||||
134 sPlysGlyAspAsnValAsnIleSerPheLysLysValLeuIleLysGlu 150
|||||
486 ACAGGGAGATATATGTAACATATCTTTCAAAAAGGCTTTAAATTAAGAA 535
|||||
151 GluAspAlaValIleTyrLysAsnGlySerPheIleHisSerValTrpAr 167
|||||
536 GAAGATGAGTATATTAATAAATATATATATATATATATATATATATAT 585
|||||
167 qHisGlyValProAspIleLeuGluVal 176
|||||
586 CATCACTACCTCATATTTTATGAGGTT 613

```

seq\_name= qb\_est1-A116662

seq\_documentation\_block:  
LOCUS A116662 718 bp mRNA linear EST 02-SEP-1998  
DEFINITION uc28493.71 533ano mouse kidney cDNA clone  
IMAGE:1481644 5' similar to qb-L06139 TYROSINE-PROTEIN KINASE  
REFSEQ:TEK PPERNSOP (HUMAN): qb-X67553 M. musculus mRNA for tek  
(MOUSE): mRNA sequence.

ACCESSION A116662

VERSION A116662.1 GI:3516986

KEYWORDS EST

SOURCE house mouse

ORGANISM

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 718)  
Marra, M., Miller, L., Allen, M., Bowles, M., Dietrich, N., Duboucq, T.,  
Geisel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M.,  
Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, R.,  
Theising, B., Wylie, T., Lennon, G., Soares, R., Wilson, R. and  
Waterston, R.

The WashU-HMI Mouse EST Project

Unpublished (1996)

Contact: Marra M/Mouse EST Project

WashU-HMI Mouse EST Project

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800

Fax: 314 286 1810  
Email: mouseest@watson.wustl.edu  
This clone is available royalty-free through LBNL; contact the  
IMAGE Consortium (info@image.lbnl.gov) for further information.  
MGI:930000

Trace considered overall poor quality

Seq primer: custom primer used

High quality sequence stop: 1.

#### FEATURES

##### source

1..718

Location/Qualifiers

/organism="Mus musculus"

/strain="C57BL"

/db\_xref="taxon:10090"

/clone\_image="1481644"

/clone\_lib="Sudano mouse kidney mklia"

/sex="female"

/dev\_stage="adult"

/lab\_host="DH10B"

/note="Organ: kidney; Vector: pME18S-FL3; Site\_1: Dral11  
(CACTGTG); Site\_2: Dral11 (CACCATGTG); 1st strand cDNA  
was primed with an oligo(dT) primer  
[ATGGCGCTTTTITTTTTT]; double-stranded cDNA was  
ligated to a Dral11 adaptor [TTTGGGTACTGG], digested  
and cloned into distinct Dral11 sites of the pME18S-FL3  
vector (5' site CACTGTG, 3' site CACCATGTG). Xho1 should  
be used to isolate the cDNA insert. Size selection was  
performed to exclude fragments <1.5kb. Library  
constructed by Dr. Sumio Sudano (University of Tokyo  
Institute of Medical Science). Custom primers for  
sequencing: 5' end primer CTTCTGCTCTAAAGCTCGG and 3' end  
primer CCACTTCAGCTCCAGACA."

HASE COUNT 190 a 164 c 185 g 176 t 3 others

ORIGIN

#### alignment\_scores:

Quality: 853.00 Length: 237

Ratio: 4.161 Gaps: 3

Percent Similarity: 86.498 Percent Identity: 79.747

#### alignment\_block:

US-09-733-764-2\_COPY\_1\_472 x A116662 ..

Align seg 1/: to: A116662 from: 1 to: 718

```

7 LeuValLeuCysGlyValSerLeuLeuLeuSerGlyThrValGly 22
|||||
6 TTAGTTCTCTGTCGAGTCAGCTTTGCTTCTTATTCGAGTACGAGG 55
|||||
22 yAlaMetAspLeuLeuLeuLeuLeuLeuLeuLeuValSerAspAlaG 39
|||||
56 GCGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATG 105
|||||
39 LuThrSerLeuThrCysIleAlaSerGlyTrpArgProHisGluProIle 55
|||||
106 AACCTTCCTTAAC.TGCTTGGCTTTGGTGGCACCACCCCAAGCCCAT 154
|||||
56 ThrIleGlyArgAspPheGluAlaLeuMetAsnGlnHisGlnAspPro 72
|||||
155 ACCATATGAGGCGGATCTGAAAGCGTTTAATGAACCCGACCCACAC 204
|||||
72 euClyValThrGlnAspValThrArgGluTrpAlaLysGlyValValTrp 88
|||||
205 TGAAGATGATATATATATATATATATATATATATATATATATAT 254
|||||
89 TysArgGlyCysAlaSerLysIleAsnGlyAlaTyrPheCysGlyArg 105
|||||
255 AAGACAGAAAAGGCGGATGAGATGAGATGAGATGAGATGAGATGAG 304
|||||
105 yValArgGlyGluAlaIleArgIleArgThrMetLysMetArgGlnGln 122
|||||
305 ACTTTCAGGATAGGCTTATTAAGGATATGAGATGAGATGAGATGAG 354
|||||

```







```

84  lys...ysValValThrpsArGlu lysAlaSerLysLysAlaSerValAla 99
   ...:|||||...|||||...|||||...|||||...|||||...
374 AACACAAGTTCTTGAAGAGAGAAAAGAGGCGAGTAAATATATCGTGT 423
100 TyrPheCysGluGlyArgValAlaGlyGluAlaIleArgIleArgThr 115
   ...:|||||...|||||...|||||...|||||...|||||...
424 TATTTCTGTGAAGTTCAGTTCGAGGAGCAAGGCTATTAAGCATACGCAAC 473
116 MetLysMetArgGlnGlnAlaSerPheLeuProAlaThrLeuThrMet 131
   ...:|||||...|||||...|||||...|||||...|||||...
474 ATGAAGATGCGTCCACACAGCACTCTTCCCTACCTACTTAACTAATG 523
132 ThrValAspLysCysAspAsnValAsnIleSerPheLysValIle 148
   ...:|||||...|||||...|||||...|||||...|||||...
524 ACCGGGAGACAGAGAGATATATGTGTATATTTTAAACAGGTGTGA 573
148 IleLysGluGluAspAlaValIleTyrLysAsnGlySerPheIleHisSer 164
   ...:|||||...|||||...|||||...|||||...|||||...
574 TIACCGAACAC 608
165 ValProArgHisGluValProAspIleLeuGluValHisLeuProHisAl 181
   ...:|||||...|||||...|||||...|||||...|||||...
609 ACAATGCGCTAC 620
181 aGlnProGlnAspAlaGlyValTyrSerAlaArgTyrIleCysCysAsn 198
   ...:|||||...|||||...|||||...|||||...|||||...
621 ... 654
198 eu...:|||||...|||||...|||||...|||||...|||||...
655 TGAAGCTTACCTGCGAGTATGCGAGAAACCTTAAAG 695
208 ValArgArgCysGlnAlaGlnLysTyrIleGlyProGluCysAsn 221
   ...:|||||...|||||...|||||...|||||...|||||...
696 ... 710

```

seq\_name: gb\_est2:BF979359

seq\_documentation\_block:

LOCUS BF979359 850 bp mRNA linear EST 23-JAN-2001  
602288449P1 NIH\_MGC\_97 Homo sapiens cDNA clone IMAGE:4374037 5',  
mRNA sequence.

ACCESSION BF979359

VERSION BF979359.1 GI:12346574

KEYWORDS EST.

SOURCE human.

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

NIH-MGC nrip://mgi.cni.nih.gov/

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: rstra@ramail.nih.gov

Tissue Procurement: Miklos Palkovits, M.D., Ph.D.

cDNA Library Preparation: Michael J. Brownstein (NIH/RI), Shiraki

Toshiyuki and Piero Carninci (RIKEN)

cDNA Library Arrayed by: The I.M.A.C.F. Consortium (I.M.C.)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LLNL10037 row: a column: 14

High quality sequence stop: 656.

location/Qualifiers

1..850

/organism="Homo sapiens"

/db\_xref="taxon.9606"

/clone="IMAGE:4374037"

/clone\_lib="NIH\_MGC\_97"

/lab\_host="PH10H"

FEATURES

source

/note="Organ: testis; Vector: phluescriptk (modified phluescript KS+); Site:1: BamHI; Site:2: SalI-XhoI (atcgag); Oligo-dT primed using primer 5'-TTTTTTTTTTTTT-3', size selected for average insert size 2.2 kb and normalized to 500 5'. This is a primary library enriched for full-length clones and constructed using the Cap-trapper method (Carninci, in preparation). Library constructed by M. Brownstein (NIH/NIH/NIH, National Institutes of Health). Note: this is a NIH\_MGC Library."

BASE COUNT 150 a 248 c 273 g 179 t

ORIGIN

alignment\_scores:

Quality: 536.00 Length: 254

Ratio: 3.045 Gaps: 7

Percent Similarity: 69.291 Percent Identity: 42.520

alignment\_block:

US-09-733-764-2\_COPY\_1\_472 x BF979359 ..

Align seq 1/1 to: BF979359 from: 1 to: 850

```

100 TyrPheCysGluGlyArgValArgCysGluAlaIleArgIleArgThrMe 116
   ...:|||||...|||||...|||||...|||||...|||||...
   9 TTCTGCTGCGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 58
116 tLysMetArgGlnGlnAlaSerPheLeuProAlaThrLeuThrMetThr 133
   ...:|||||...|||||...|||||...|||||...|||||...
   59 GCATCAACAGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 108
133 aLaspLysCysAspAsnValAsnIleSerPheLysValIleLys 149
   ...:|||||...|||||...|||||...|||||...|||||...
104 TCAATCAAGAGTGAATCAATCAATCAATCAATCAATCAATCAATCA 158
150 GluGluAspAlaValIleTyrLysAsnGlySerPheIleHisSerVal 166
   ...:|||||...|||||...|||||...|||||...|||||...
159 CAGACACAGCGGAGATGCGAGATGCGAGATGCGAGATGCGAGATGCG 208
166 aArgHisGlnValProAsp 116LeuGluValHisLeuProH 180
   ...:|||||...|||||...|||||...|||||...|||||...
204 CAGATCAACAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 252
180 IsAlaGlnProGlnAspAlaGlyValTyrSerAlaArgTyrIleGly 196
   ...:|||||...|||||...|||||...|||||...|||||...
253 AITGCGAGCGATGATGAGCGGATGATGAGCGGATGATGAGCGGATG 402
197 AsnLeuPheThrSerAlaPheThrArgLeuIleValArgArgCysGlu 213
   ...:|||||...|||||...|||||...|||||...|||||...
303 AGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 352
213 aGlnLysTrpGlyProGluCysAsnHisLeuCysThrAlaCysMetAsn 230
   ...:|||||...|||||...|||||...|||||...|||||...
353 TGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 402
230 snCysPheCysHisCysAspThrGlyGlyCysIleCysThrGlyPhe 246
   ...:|||||...|||||...|||||...|||||...|||||...
404 GAGATGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATG 452
247 MetGlyArgThrCysGluLysAlaCysGlnHisThrPheGlyArg 253
   ...:|||||...|||||...|||||...|||||...|||||...
453 ACTGCGAGCGGATGAGATGAGATGAGATGAGATGAGATGAGATG 502
253 rCysLysGluArgCysSerGlyGlnGluGlyCysLysSerTyrValPhe 280
   ...:|||||...|||||...|||||...|||||...|||||...
503 CTGCGAGCGGATGAGATGAGATGAGATGAGATGAGATGAGATG 551
280 yLysProAspProTyrCysSerCysAlaThrGlyThrLysCysGly 296
   ...:|||||...|||||...|||||...|||||...|||||...
552 CTGCGAGCGGATGAGATGAGATGAGATGAGATGAGATGAGATG 599
296 uCysAsnGluAlaCysHisProGlyPheTyrGlyProAspCysLys 313
   ...:|||||...|||||...|||||...|||||...|||||...

```









```

Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshingenes.nig.ac.jp.
Location/Qualifiers
  1..625
Source
  Zoonosis="Xenopus laevis"
  /db_xref="taxon:8355"
  /clone="XL054122"
  /clone_lib="NIBH Morchii normalized Xenopus tailbud
  library"
  /tissue_type="whole embryo"
  /dev_stage="stage 25"
  195 a 124 c 138 q 168 L
BASE COUNT
ORIGIN

alignment_scores:
  Quality: 427.00      Length: 171
  Ratio: 3.436        Gaps: 2
  Percent Similarity: 74.854 Percent Identity: 46.744

alignment_block:
  US-09-733-764-2_COPY_1_472 x BJ066014
  Align seq 1/1 to: BJ066014 from: 1 to: 625

301 AlAcGSHSHstrGlyPheTrpCysGlyAspCysGlySerGly 317
|||||
49 GCTTCTCAGATGCTTTTATGACCGGAGTGCATAATTCAGATGCTTTC 98
|||||
317 sAsAsAGlyGluM-CysAspArPheGlyGlyCysLeu-CysSerPro 334
|||||
99 TAACAAAGGA...ACTGTCAATACATTTAAGGCGTCCATTTGCCACAAAC 145
|||||
334 lyTrpGlnGlyGluGlnCysGluArgGluGlylleProArgMetThrPro 350
|||||
146 GCTGCGAGAGTACATTGTGAATATTAAGGATGCTTAAATATACATCA 195
|||||
351 LysIleValAspLeuProAspPheIleGluValAsnSerGly...LysPhe 366
|||||
196 ACAATTTTGGAAACCAACAAACATTAGAGCTTAATTTTACCTGCGCAAAAT 245
|||||
366 eAsnProIleCysIysAlaSerGlyTrpProLeuProThrAsnGluGluM 383
|||||
245 CACCTTCTACTTCTAAACCTCACTTCAACCTTCTTCTTCTTCTTCTTCT 295
|||||
383 etThrLeuValIysProAspGlyThrValLeuHisProIysAspPheAsn 399
|||||
296 TCAACTATTATTAACAAAGATGAAACATTGTTTGTGTGTGTGTGTGT 445
|||||
400 HisThrAspHisPheSerValAlaIlePheThrIleHisArgIleLeuP 416
|||||
346 ACCAAATATCAGAGGAGTATGCTGCTTTTAAAGGCGGCAACAAATTAAG 395
|||||
416 oProAspSerGlyValTrpValCysSerValAsnThrValAlaGlyMetV 433
|||||
396 AAGAGGCTTAAGTATTTGGATATGCTAGTGTCAAGACAGATAGCGGCA 445
|||||
433 alGluIysProPheAsnIleSerValIysValLeuProIysProLeuAsn 449
|||||
446 TTGAGAGCGCTTTCTATATTTTCTTAAAGCTCCACCACTGCTTAAATAC 495
|||||
450 AlaProAsnValIleAspThrGlyHisAsnProAlaValIleAsnIle 466
|||||
496 ATGCTAAAGCTAAATACAGCTGCGCTTCAATATCTTACTATTGAGGTGAA 545
|||||
466 rSerGluProTyr 470
|||||
546 TGCATATCCCTAC 558

seq_name: gb_est2:BF850180
seq_documentation_block:

```

```

LOCUS      BF850180                299 bp      mRNA      linear      EST 16-JAN-2001
DEFINITION IL5-EN0084-171100-248-q05 EN0084 Homo sapiens cDNA, mRNA sequence.
ACCESSION  BF850180
VERSION    BF850180.1   GI:12237342
KEYWORDS   EST
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
REFERENCE  1 (bases 1 to 299)
AUTHORS   Dias Neto,F., Garcia Correa,R., Vertovski-Almeida,S., Briones,M.R.,
            Razaai,M.A., da Silva,W Jr., Zagari,A., Rordin,S., Costa,F.P.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bala,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jondeneel,C.V., O'Hare
            ,M.J., Soares,F., Brentani,R.R., Reis,I.F., de Souza,S.J. and
            Simpson,A.J.
TITLE      Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
JOURNAL    Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE    20020663
COMMENT    Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            http://www.cude-qibq.ql/sciapes/qibq.milz.pl/IL5-IL5A2-IL5-EN0084-
            171100-248-q05q3-2000-11-17kl4-1)
            Seq primer: puc 18 forward
            High quality sequence stop: 298.
FEATURES   Location/Qualifiers
            1..299
             /organism="Homo sapiens"
             /db_xref="taxon:9606"
             /clone_lib="EN0084"
             /dev_stage="Adult"
             /note="organ. lung_normal; Vector: puc18; Site:1; Small;
             Site:2; Small; A mini-library was made by cloning products
             derived from opfetes pop (n.s. Letiers patent application
             No. 196.716 - Ludwig Institute for Cancer Research)
             Profiles into the puc 18 vector. Reverse transcription of
             tissue mRNA and cDNA amplification were performed under
             low stringency conditions."
BASE COUNT 52 a 90 c 102 g 55 t
ORIGIN

alignment_scores:
  Quality: 345.00      Length: 92
  Ratio: 4.726        Gaps: 0
  Percent Similarity: 79.348 Percent Identity: 55.435

alignment_block:
  US-09-733-764-2_COPY_1_472 x BF850180
  Align seq 1/1 to: BF850180 from: 1 to: 299

209 ArgArgCysGlnAlaGlnLysTrpGlyProGluCysAsnHisLeuCysTh 225
|||||
1 GGGGATGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 50
|||||
225 rAlaCysMetAspAsnGlyValCysHisGluAspThrGlyGlyCysIleC 242
|||||
51 AGGTGAGTAAATATGAGAGTGTGTATTAAGACATGACAGAGAAATGATG 100
|||||
242 ystProGlyPheMetGlyArgThrCysGluCysAlaCysIleLeuLeu 258
|||||
101 GGGGATGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 150
|||||

```

